

# Entwickler von komplexen Mensch-Maschine-Systemen: Analyse des Einflusses von deren Leistungsdeterminanten auf den Entwicklungsprozess und das -ergebnis

Dissertation  
zur Erlangung des akademischen Grades  
doctor rerum naturalium (Dr. rer. nat.)  
im Fach Psychologie

eingereicht an der Lebenswissenschaftlichen Fakultät  
der Humboldt-Universität zu Berlin

von Dipl.-Psych. Saskia Kain

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät:

Prof. Dr. Bernhard Grimm

Gutachter:

1. Prof. Dr. Hartmut Wandke (Humboldt-Universität zu Berlin)
2. Prof. Dr.-Ing. Matthias Rötting (Technische Universität Berlin)
3. Prof. Dr. Pierre Sachse (Universität Innsbruck)

Eingereicht am: 15.01.2016

Tag der Verteidigung: 15.09.2016

## **Abstract**

Automation is currently to be found in all aspects of life. “Reliable” automation is believed to take the pressure off, to compensate, or even displace the “unreliable” human (user/operator). However, automation can be fraught with problems. Among other things, it can appear that individuals experience a loss of situational awareness. Moreover, as a consequence of being developed by “unreliable” humans (developers) automations can be susceptible to error. In these instances, the user/operator often has to restore the normal state of a system within a very short time.

As a response, between 2004-2013 Project ATEO ArbeitsTeilung Entwickler Operateur (division of work between developers and operators) pursues an, so far in researches only theoretical considered, approach to optimize the function allocation between human and machine. It is about an optimization of the division of responsibilities between operators and developers of complex, dynamic, technical systems including automations.

This thesis concerns the performance determinants of these developers. Literature regarding the previous findings of performance determinants research were, together with Rocholl (2010), gathered into a descriptive "model of developers' performance determinants" and validated in an expert-workshop. For an initial experiment the test environment "Socially Augmented Microworld" (SAM) was used as system. The complexity of this system was prior heightened to ensure that the conception of automatic functions for this system was not too easy for professional developers. The enhancements were determined experimentally in four preliminary studies. With the use of an experiment a comparison of the new with the old test environments it was proved that the efforts were successful. Moreover, an automations software was designed to enable the implementation and configuration of the developer-created automations. Subsequently, the potential performance determinant "information availability" was examined experimentally by using 90 professional developers. To capture the dependent variable "quality of automation concepts" an assessment tool was developed. The results of the experiment with the professional developers showed that the control variables “experience index”, “number of jokes” and “achievement motive hope of success” had a higher influence on the quality of the design result as the systematically varied "information availability". Only by eliminating the effect of these control variables a significant influence of the performance determinant "information availability" on the quality of automation conceptions was attested.

## **Zusammenfassung**

Automation ist heute in allen Lebensbereichen zu finden. Die „zuverlässige“ Automation soll den „unzuverlässigen“ Menschen (Benutzer/Operateur) entlasten, kompensieren, teilweise sogar ersetzen. Das kann jedoch problematisch sein. Beim Menschen kann sich u. a. ein Verlust des Situationsbewusstseins zeigen. Eine Automation kann darüber hinaus Fehler machen, wurden sie doch von „unzuverlässigen“ Menschen (Entwicklern) entwickelt. In einem solchen Fall muss der Benutzer/Operateur oftmals innerhalb kürzester Zeit den Normalzustand eines Systems wiederherstellen.

Deshalb verfolgte das in den Jahren 2004 bis 2013 durchgeführte Projekt ATEO (ArbeitsTeilung Entwickler-Operateur) einen in der Forschung bisher nur theoretisch betrachteten Ansatz zur Optimierung der Funktionallokation zwischen Mensch und Maschine. Hierbei handelt es sich um die Optimierung der Arbeitsteilung zwischen Operateuren und Entwicklern von komplexen, dynamischen, technischen Systemen inklusive Automationen.

Diese Dissertation beschäftigt sich mit den Leistungs determinanten dieser Entwickler. Die bisherigen in der Literatur beschriebenen Befunde zu diesen Leistungs determinanten wurden gemeinsam mit Rocholl (2010) zu einem deskriptiven „Modell der Leistungs determinanten von Entwicklern“ zusammengetragen und dieses in einem Expertenworkshop validiert. Für eine erste experimentelle Prüfung wurde die Versuchsumgebung „Socially Augmented Microworld“ (SAM) als System verwendet. Deren Komplexität wurde zuvor erhöht, damit sich für professionelle Entwickler die Konzeption von Automaten für dieses System nicht zu leicht gestaltet. Die Erweiterungen wurden in vier Voruntersuchungen experimentell ermittelt. Durch einen experimentellen Vergleich der neuen mit der alten Versuchsumgebung konnte nachgewiesen werden, dass die Bemühungen erfolgreich waren. Zusätzlich wurde eine Automatensoftware zur Implementierung und Konfiguration der konzipierten Automaten entwickelt. Anschließend wurde die potentielle Leistungs determinante „Informationsverfügbarkeit“ unter Einsatz von 90 professionellen Entwicklern experimentell beleuchtet. Zur Erfassung der abhängigen Variable „Güte der Automatenkonzepte“ wurde eigens ein Bewertungsinstrument entwickelt. Durch die Untersuchung konnte gezeigt werden, dass die Kontrollvariablen „Erfahrungsindex“, „Witzanzahl“ und „Leistungsmotiv HE“ einen höheren Einfluss auf die Güte des Konzeptionsergebnisses haben als die systematisch variierte „Informationsverfügbarkeit“. Erst durch Eliminierung der Wirkung dieser Kontrollvariablen konnte ein signifikanter Einfluss der zur Verfügung gestellten Informationen auf die Güte der Automatenkonzepte nachgewiesen werden.

### **Keywords**

function allocation, operator, developer, assistance system, automation, supervisory control, performance determinants, Socially Augmented Microworld (SAM)

### **Schlagwörter**

Funktionsteilung, Arbeitsteilung, Operateur, Entwickler, Automation, Assistenzsystem, Automatik, Prozessüberwachung und -führung, Leistungsdeterminante, sozial angereicherte Mikrowelt (SAM)

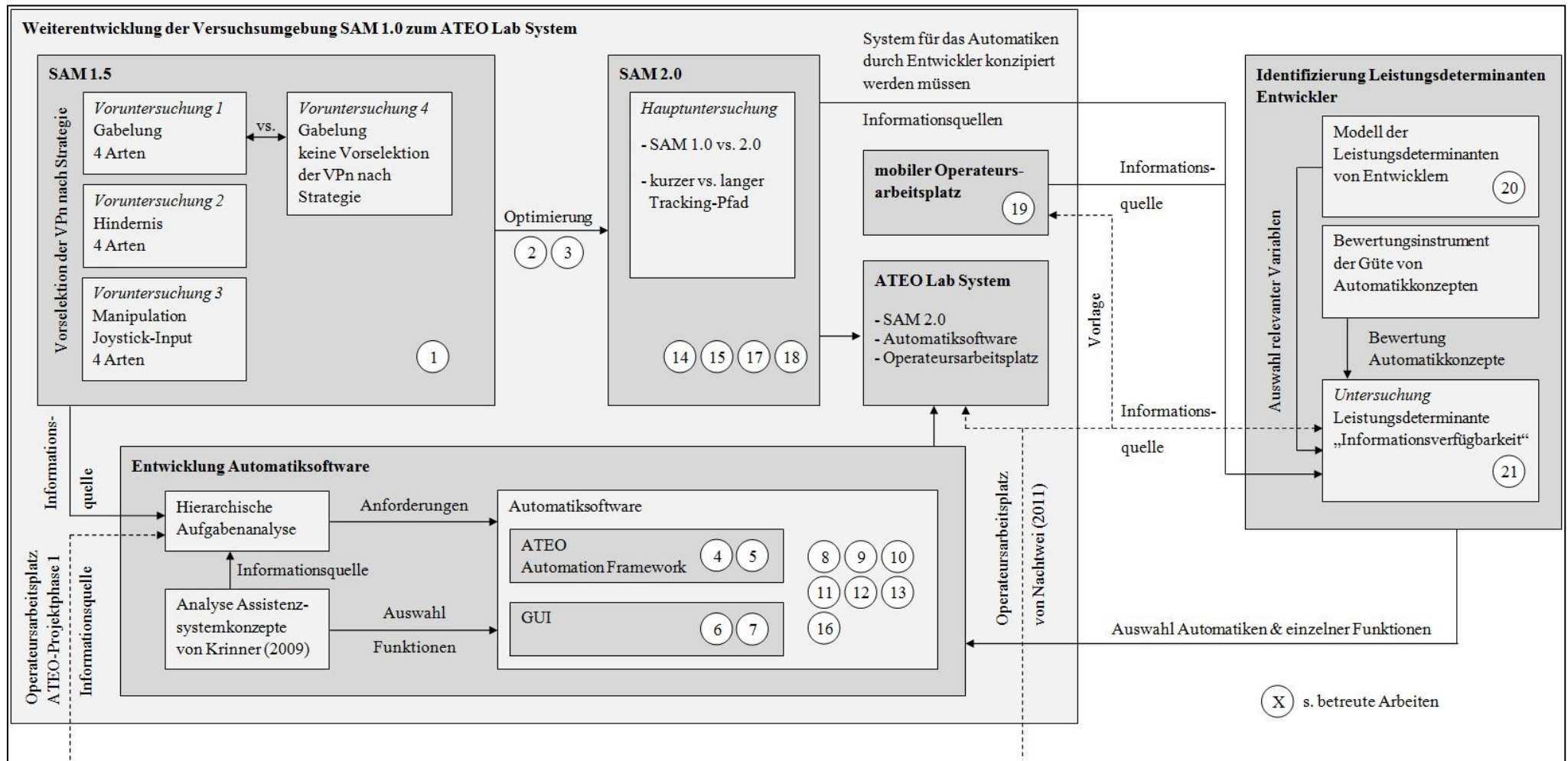


Abbildung 1: Darstellung der einzelnen Arbeitspakete dieser Dissertation und ihrer Zusammenhänge

*Anmerkungen.* Für die Bewältigung einzelner Arbeitspakete relevante Dokumente oder Softwarekomponenten von anderen ATEO-Projektmitgliedern werden von außen über eine gestrichelte Linie in den Arbeitspaketeplan eingeführt. Ein Kreis mit Nummer kennzeichnet eine durch die Autorin betreute Studentenarbeit (s. Betreute Arbeiten) im Rahmen eines Arbeitspakets.

## **Betreute Arbeiten**

### **Weiterentwicklung der Versuchsumgebung SAM 1.0 zum ATEO Lab System**

#### ***Von SAM 1.0 zu SAM 1.5***

1. Niestroj, N. (2008). Erweiterung des ATEO-Systems zur Komplexitätserhöhung von SAM. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***Von SAM 1.5 zu SAM 2.0***

2. Hildebrandt, M. (2008). Analyse und Evaluierung der Architektur des ATEO-Systems. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
3. Niestroj, N. (2009). Vernetzung im ATEO Projekt aus inhaltlicher und technischer Sicht. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

### ***Entwicklung einer Automatiksoftware***

#### ***ATEO Automation Framework***

4. Kesselring, K. (2009). Entwicklung einer Softwarekomponente zur Systemüberwachung und -führung in einer psychologischen Versuchsumgebung. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
5. Hasselmann, M. (n. D.). Erweiterung einer Softwarekomponente für Automaten zur Systemprozessüberwachung und -führung als Bestandteil einer psychologischen Versuchsumgebung. Unpublished Diplomarbeit in Vorbereitung. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***GUI für Automatiksoftware***

6. Fuhrmann, E. (2010). Entwicklung eines GUI für die Konfiguration der Software-Komponente zur Systemprozessüberwachung und -kontrolle in einer psychologischen Versuchsumgebung. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
7. Kosjar, N. (2011). Die Gebrauchstauglichkeit des Automaten-GUI im Projekt Arbeitsteilung Entwickler Operateur (ATEO). Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***Implementierung von Automatikfunktionen***

8. Kosjar, N. (2012). Ein Ereignis-System für das ATEO Automation Framework sowie die Implementierung und Testung von auditiven und visuellen Hinweisen. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
9. Seid, A. (2012). Entwicklung und Konfiguration von visuellen Anzeigen für die Objektsteuerung im Projekt Arbeitsteilung Entwickler-Operateur (ATEO). Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

10. Weidner-Kim, H. (2014). Konzeption und Implementation einer optimalen Rückführung in ATEO SAM. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
11. Weidner-Kim, H. (2015). Konzeption und Implementation einer verbesserten optimalen Rückführung in ATEO SAM. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
12. Wickert, A. (2012). Fehler und deren Vermeidung bei der Programmierung und Konfiguration von Automaten im ATEO-Projekt. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
13. Wickert, A. (2013). Klassifizierung und Entwicklung von Automaten für Eingriffe in der Socially Augmented Microworld (SAM). Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***Logfile-Analyse-Tool LFA 1.0 bis 2.0***

14. Hampel, T. (2012). Konzeption und Implementation eines Analysetools zur Auswertung von Logfiles einer komplexen, dynamischen Versuchsumgebung. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
15. Hampel, T. (2012). Weiterentwicklung und Verbesserung einer Analysesoftware für dynamische Versuchsumgebungen. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
16. Seid, A. (2012). Erweitern der Log-Datei und der Analyse von Log-Dateien im ATEO Projekt. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***SAM Performance***

17. Dullat, S. (n. D.). Optimierungsmöglichkeiten der Performance einer softwarebasierten Versuchsumgebung. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***SAM<sub>Java</sub>***

18. Hildebrandt, M. (2009). Reengineering der Architektur der SAM-Komponente des ATEO-Systems und deren Prototypische Implementation in Java. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

#### ***Mobiler Operateursarbeitsplatz***

19. Hasselmann, M. (2010). ATEO: Der Operateursarbeitsplatz in mobiler Umgebung. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

---

**Identifizierung der Leistungsdeterminanten von Entwicklern*****Modell der Leistungsdeterminanten von Entwicklern***

20. Rocholl, A. (2010). Determinanten der Leistung von Entwicklern von Automaten und Assistenzsystemen zur Überwachung und Führung dynamischer technischer Prozesse. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Psychologie.

***Videodatenauswertung Untersuchung Leistungsdeterminante „Informationsverfügbarkeit“***

21. Dodt, K. (2012). Anwendung des IKD (Instrument zur Kodierung von Diskussionen) zur Ermittlung des Einflusses von sozialen Prozessen auf den Konzeptionsprozess und das Ergebnis. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Psychologie.



## **Danksagung**

Mein ausdrücklicher Dank gilt meinem Doktorvater, Herrn Prof. Dr. Hartmut Wandke, für seine engagierte Betreuung, die anregenden und motivierenden Gespräche, die vielen schnellen und wertvollen Unterstützungen und Vorschläge. Ferner sei ihm gedankt für die Mitbetreuung der vielen Studentenarbeiten im Rahmen meiner Dissertation.

Ebenso gilt ein sehr herzlicher Dank Herrn Prof. Dr.-Ing. Matthias Rötting für die Zweitbetreuung dieser Arbeit, seine Unterstützung und Bereitschaft, auch im Feierabend Fragen zu beantworten.

Ein besonderer Dank gilt Herrn Prof. Dr. Pierre Sachse für die Übernahme der Gutachtertätigkeit.

Der Deutschen Forschungsgemeinschaft (DFG) möchte ich für die Förderung meiner Forschungstätigkeit danken.

Herrn Prof. Dr. Klaus Bothe (HU Berlin) danke ich vielmals für die Mitbetreuung der Arbeiten der Informatikstudenten im Rahmen meiner Dissertation und für die Möglichkeit, den Inhalt eines seiner Seminare mitzugestalten. Nur dadurch konnte eine Vielzahl interessierter Studenten der Informatik zur Weiterentwicklung der Versuchsumgebung gewonnen werden.

Ein ganz besonderer Dank gebührt meinem Partner, meinem Bruder und meinen Eltern, die mir allen notwendigen Freiraum gaben, immer Verständnis zeigten und mich zu jeder Zeit unterstützten.

Für das Korrekturlesen des englischen Abstracts bin ich meinem Schwager und für das Korrekturlesen einzelner Kapitel bin ich Ariane Keitel, Kathrin Dodt, Yvonne Rudelt und Sebastian Graichen sehr dankbar. Ein besonderer Dank gilt hierbei Sarah Pleyer und meiner Familie, welche nahezu die gesamte Dissertationsschrift gesichtet haben.

Nicole Scheibner danke ich für die Bekräftigung meines statistischen Vorgehens.

Ich bedanke mich bei allen VPn, ohne deren Bereitschaft und Engagement ich mein Dissertationsvorhaben nicht umsetzen hätte können. Ebenso danke ich sämtlichen Unternehmen und Instituten, welche mir Entwickler bereitgestellt haben. In diesem Zusammenhang gilt ein besonderer Dank Herrn Henning Butz für die Unterstützung bei der Akquise von professionellen Entwicklern als VPn. Ohne seine Kontakte wären mir viele Türen verschlossen geblieben und eine Untersuchung mit professionellen Entwicklern kaum möglich gewesen.

Bei Barbara Gross möchte ich mich für die Bereitstellung der objektiven Tracking-Daten aus ihrer Untersuchung von SAM 1.0 bedanken.

Zu guter Letzt gilt ein ganz besonderes Dankeschön meinem Projektpartner Jens Nachtwei und allen Studenten, die unter „Betreute Arbeiten“ namentlich aufgelistet sind. Ohne Euch hätte diese Arbeit nicht realisiert werden können. Es war eine spannende und lehrreiche Zeit mit Euch.

# Inhaltsverzeichnis

<b>Tabellenverzeichnis .....</b>	<b>XII</b>
<b>Abbildungsverzeichnis .....</b>	<b>XV</b>
<b>Abkürzungsverzeichnis.....</b>	<b>XVII</b>
<b>1. Einleitung .....</b>	<b>1</b>
<b>1.1 Einführung in das Forschungsthema und Ableitung der Ziele .....</b>	<b>1</b>
<b>1.2 Aufbau des Projekts ATEO .....</b>	<b>4</b>
<b>1.3 Aufbau der Arbeit.....</b>	<b>6</b>
<b>PART I – Weiterentwicklung einer Versuchsumgebung .....</b>	<b>8</b>
<b>2. ATEO-Projektphase 1: Entwicklung einer Versuchsumgebung .....</b>	<b>9</b>
<b>2.1 Ausgangslage .....</b>	<b>9</b>
2.1.1 Klassische Mikrowelt.....	9
2.1.1.1 Determinismus vs. Stochastik .....	10
<b>2.2 Entwicklung einer Mikrowelt als ATEO-Versuchsumgebung.....</b>	<b>11</b>
2.2.1 Squeak als Programmierungsumgebung.....	12
2.2.2 Versuchsumgebung SAM 1.0 .....	12
2.2.2.1 Komplexität von SAM 1.0 .....	15
2.2.2.2 Mängel an SAM 1.0 .....	17
<b>3. ATEO-Projektphase 2: Weiterentwicklung der Versuchsumgebung .....</b>	<b>19</b>
<b>3.1 Methodik der Weiterentwicklung von SAM .....</b>	<b>19</b>
3.1.1 Steigerung der Leistungsmotivation der MWB .....	19
3.1.1.1 Leistungsabhängige Entlohnung .....	20
3.1.1.2 Leistungsfeedback.....	20
3.1.2 Erhöhung der Glaubhaftigkeit der „Teamaufgabe“ .....	21
3.1.3 Mehr Entscheidungssituationen in Bezug auf Geschwindigkeit vs. Genauigkeit .....	23
3.1.3.1 Gabelungen.....	23
3.1.3.2 Hindernisse.....	24
3.1.4 Manipulation des Joystick-Inputs.....	26
<b>3.2 Methodik der experimentellen Überprüfung der Weiterentwicklung.....</b>	<b>26</b>
3.2.1 Voruntersuchungen: Neue SAM-Komponenten .....	27
3.2.1.1 Fragestellung .....	27
3.2.1.2 Methoden.....	28
3.2.1.3 Ergebnisse .....	40

3.2.1.4 Diskussion und Interpretation der Ergebnisse zur Forschungsfragestellung 1 .....	46
3.2.1.5 Fazit und Kritik .....	47
3.2.2 Voruntersuchung: Vorselektion der MWB .....	48
3.2.2.1 Fragestellung und Hypothesen .....	48
3.2.2.2 Methoden.....	51
3.2.2.3 Ergebnisse .....	52
3.2.2.4 Interpretation und Diskussion der Ergebnisse zur Forschungsfragestellung 2 .....	54
3.2.2.5 Fazit, Kritik und Ausblick.....	56
3.2.3 Hauptuntersuchung: Vergleich SAM 1.0 vs. 2.0 .....	58
3.2.3.1 Fragestellungen und Hypothesen .....	58
3.2.3.2 Methoden.....	64
3.2.3.3 Ergebnisse .....	74
3.2.3.4 Interpretation und Diskussion der Ergebnisse.....	93
3.2.3.5 Kritik .....	102
3.2.3.6 Fazit und Ausblick .....	103
<b>3.3 Von SAM 2.0 zum ATEO Lab System.....</b>	<b>105</b>
3.3.1 Entwicklung einer Automatiksoftware für SAM .....	105
3.3.1.1 Hierarchische Aufgabenanalyse für eine Automatiksoftware.....	105
3.3.1.2 Konzeption und Implementierung einer Automatiksoftware.....	106
3.3.1.3 Graphical User Interface der Automatiksoftware .....	108
3.3.2 Entwicklung eines Operatorsarbeitsplatzes für SAM.....	113
3.3.2.1 Entwicklung des Operatorsarbeitsplatzes für 13,3“ Monitor .....	114
3.3.3 ATEO Lab System (ALS).....	115
<b>3.4 Fazit zur Umsetzung der Komponenten des ATEO Lab Systems in Squeak .....</b>	<b>116</b>
<b>PART II – Determinanten von Entwicklern.....</b>	<b>118</b>
<b>4. Theoretischer Hintergrund.....</b>	<b>119</b>
<b>4.1 Automation in Mensch-Maschine-Systemen .....</b>	<b>119</b>
4.1.1 Begriffsbestimmung „Automatisierung“, „Automation“ und „Funktion“ .....	119
4.1.2 Probleme durch Automation .....	121
4.1.2.1 Unangemessenes Vertrauen in die Automation .....	121
4.1.2.2 Sinkendes Situationsbewusstsein .....	122
4.1.2.3 Fertigkeitsverlust.....	122
4.1.3 Mensch-Maschine-Funktionsteilung.....	123
4.1.4 Begriffsbestimmung „Automatik“ und „Assistenzsystem“ .....	125
4.1.5 Automatisierungsstufen.....	126
4.1.6 Optimales Ausmaß der Automation.....	130
4.1.6.1 Flexible Automation.....	133

---

<b>4.2 Entwicklung von Mensch-Maschine-Systemen.....</b>	<b>134</b>
4.2.1 Entwickler von technischen Systemen.....	135
4.2.1.1 Vorgehensmerkmale erfolgreicher Entwickler .....	136
4.2.2 Merkmale der Entwicklungsaufgabe.....	138
4.2.2.1 Design Problem Solving.....	140
4.2.3 Entwicklungsprozess.....	142
4.2.3.1 Systematisches vs. opportunistisches Vorgehen bei der Entwicklung.....	144
4.2.3.2 Frühe Phasen des Entwicklungsprozesses.....	145
<b>5. Determinanten der Leistung von Entwicklern .....</b>	<b>147</b>
<b>5.1 ATEO-Projektphase I: Identifizierung der Leistungsdeterminanten .....</b>	<b>147</b>
5.1.1 Untersuchung Leistungsdeterminante „Art der Information über das System“ ....	147
5.1.2 Untersuchung „Ressourcenmaximierung“ .....	148
5.1.3 Expertenevaluation der Assistenzsystemkonzepte.....	149
5.1.4 Mängel bei der Identifizierung der Leistungsdeterminanten .....	150
<b>5.2 ATEO Projektphase 2: Identifizierung der Leistungsdeterminanten .....</b>	<b>151</b>
5.2.1 Modell der Leistungsdeterminanten von Entwicklern .....	151
5.2.1.1 Modell Output-Operationalisierung „Teamleistung“ .....	154
5.2.2 Bewertungsinstrument der Güte von Automatikkonzepten .....	156
5.2.2.1 Bewertungsfunktionsmatrix .....	159
5.2.2.2 Bewertungsfragebogen.....	177
5.2.3 Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“ .....	180
5.2.3.1 Fragestellung und Hypothesen .....	180
5.2.3.2 Methoden.....	184
5.2.3.3 Ergebnisse .....	204
5.2.3.4 Interpretation und Diskussion der Ergebnisse.....	218
5.2.3.5 Kritik .....	228
5.2.3.6 Ausblick .....	230
<b>6. Gesamtfazit und Ausblick.....</b>	<b>232</b>
<b>Literaturverzeichnis.....</b>	<b>236</b>
<b>Anhang liegt gesondert vor.</b>	

## Tabellenverzeichnis

Tabelle 1:	Darstellung des jeweils aktiven MWBs in den 16 Tracking-Durchgängen eines Versuchsdurchlaufs .....	33
Tabelle 2:	Variantenreihenfolgen der uV „Gabelung“ .....	34
Tabelle 3:	Variantenreihenfolgen der uV „statisches Hindernis“ .....	34
Tabelle 4:	Variantenreihenfolgen der uV „Manipulation des Joystick-Inputs“ .....	34
Tabelle 5:	Reihenfolgen der Manipulation des Joystick-Inputs der MWB .....	34
Tabelle 6:	Mittelwerte und Standardabweichungen der Differenzen bei der horizontalen und vertikalen Joystick-Auslenkung der MWB, der benötigten Zeit und der Fehler beim Objekt-Tracking im Bereich vor der Gabelung.....	42
Tabelle 7:	Häufigkeiten und Prozente der Gabelungspfadwahl sowie Mittelwerte und Standardabweichungen der benötigten Zeit und der Fehler beim Objekt-Tracking im Bereich innerhalb der Gabelung .....	43
Tabelle 8:	Mittelwerte und Standardabweichungen der Differenzen bei der horizontalen und vertikalen Joystick-Auslenkung der MWB, der benötigten Zeit und der Fehler beim Objekt-Tracking beim Einsatz von Hindernissen .....	44
Tabelle 9:	Mittelwerte und Standardabweichungen der Differenzen bei der horizontalen und vertikalen Joystick-Bewegung der MWB, der benötigten Zeit und Fehler beim Objekt-Tracking beim Einsatz der Joystick-Manipulation.....	46
Tabelle 10:	Die, für die Datenanalyse, relevanten erhobenen Variablen respektive eingesetzten Fragebogen, die verwendeten Antwortskalen und Messzeitpunkte in der Untersuchung von SAM 2.0.....	68
Tabelle 11:	Reihenfolge der Gabelungen und Hindernisse in den kooperativen Tracking-Durchgängen .....	72
Tabelle 12:	Mittelwert und Standardabweichung des Flächenfehlers beim Tracking in Millionen Pixel pro Durchgang und über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0.....	79
Tabelle 13:	Mittelwert und Standardabweichung der benötigten Zeit zum Tracking in Sekunden pro Durchgang und über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0.....	80
Tabelle 14:	Mittelwert, Standardabweichung und Standardfehler des Mittelwerts des Flächenfehlers beim Tracking in Millionen Pixel für den kurzen und langen Tracking-Pfad.....	83

Tabelle 15:	Deskriptive Statistiken der Beurteilungen durch die MWB bzgl. Instruktionsverständlichkeit, Instruktionsbefolgung und Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen.....	85
Tabelle 16:	Mittelwert und Standardabweichung der Selbsturteile der MWB worauf beim Tracking mehr geachtet wurde je Messzeitpunkt und über alle Messzeitpunkte hinweg.....	88
Tabelle 17:	Mittelwerte und Standardabweichungen der Beurteilungen der eigenen Wichtigkeit und der des anderen MWBs ein gutes Tracking-Ergebnis zu erreichen.....	90
Tabelle 18:	Deskriptive Statistiken der Beurteilung der Zufriedenheit mit der Teamleistung nach den kooperativen Tracking-Durchgängen durch die MWB.....	91
Tabelle 19:	Mittelwert und Standardabweichung der subjektiven Beanspruchung der MWB und subjektiven Komplexität der einzelnen Tracking-Durchgänge für beide MWB zusammen pro Messzeitpunkt.....	94
Tabelle 20:	Merkmale von komplexen Problemen nach Dörner (1989), deren Bedeutungen sowie äquivalente Merkmale von Entwicklungsaufgaben.....	140
Tabelle 21:	Struktur der Bewertungsfunktionsmatrix mit Darstellung der Hauptkategorie „Funktionsklasse“ sowie einzelner Kategorien von deren Unterkategorie, der Hauptkategorie „Pfadelement“ und einzelner Kategorien von deren Unterkategorie „Situation“ sowie exemplarischer Funktionen. Ein X kennzeichnet, für welche Situation eine Funktion geplant ist.....	160
Tabelle 22:	Eine exemplarische Funktion der Funktionsklasse „Informationsanalyse“ mit deren Anforderungen und gewichteten Auslegungsoptionen.....	176
Tabelle 23:	Items des „Bewertungsfragebogens für die Güte von Automatikkonzepten“ .....	177
Tabelle 24:	Item „Anzeigenkomplexität“ des „Bewertungsfragebogens Güte von Automatikkonzepten“ samt Ausprägungsextreme und Begründung .....	179
Tabelle 25:	Die, für die Datenanalyse, relevanten erhobenen Variablen respektive eingesetzten Fragebogen, die verwendeten Antwortskalen und Messzeitpunkte in der Untersuchung „Informationsverfügbarkeit“ .....	190
Tabelle 26:	Geschlechterverteilung in den Teams, Ausbildung/Hochschulabschluss und berufliche Position der VPn der drei Bedingungen der uV „Informationsverfügbarkeit“ und für die Gesamtstichprobe in Form von Häufigkeit .....	202

Tabelle 27:	Mittelwert und Standardabweichung der Teamwerte (Mittelwert der Mitgliederwerte) einiger Kontrollvariablen pro Bedingung der uV „Informationsverfügbarkeit“ sowie Ergebnisse der statistischen Analysen (Varianzanalyse, H-Test von Kruskal-Wallis, t-Test für unabhängige Stichproben und Mann-Whitney-U-Test mit Effektgrößen und Teststärken) zur Überprüfung der Gleichheit der Stichproben dieser drei Bedingungen ..... 203
Tabelle 28:	Produkt-Moment-Korrelationen nach Pearson zwischen Input-, Prozessvariablen und der Output-Variable zur Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“ mit der Output-Variable „Teamleistung“ ..... 209
Tabelle 29:	Rangkorrelationen tau von Kendall zwischen Input-Variablen und der Output-Variable zur Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“ mit der Output-Variable „Teamleistung“ ..... 210
Tabelle 30:	Schrittweise multiple lineare Regression mit den Prädiktoren „Erfahrungsindex“, „Residuum Witzanzahl“ und „Leistungsmotiv HE“ sowie dem Kriterium „Güte der Automatikkonzepte“ ..... 213
Tabelle 31:	Anzahl der Teams, die Funktionen konzipierten, Mittelwert und Standardabweichung der Funktionsanzahl und der erreichten Punktzahl der Teams je Unterkategorie der Hauptkategorie „Funktionsklasse“ der Funktionsmatrix pro Bedingung der uV „Informationsverfügbarkeit“ zur explorativen Überprüfung von Unterschieden bei den Automatikkonzepten ..... 216

## Abbildungsverzeichnis

Abbildung 1:	Darstellung der einzelnen Arbeitspakete dieser Dissertation und ihrer Zusammenhänge .....	IV
Abbildung 2:	Die drei Gabelungsarten von SAM 1.0.....	17
Abbildung 3:	Geschwindigkeitsanzeige im Objekt in Form eines sich gegen den Uhrzeigersinn auffüllenden schwarzen Kreises in SAM 1.5, von links null Prozent (Stillstand) bis rechts 100 Prozent Geschwindigkeit.....	21
Abbildung 4:	Gabelungen von SAM 1.5.....	24
Abbildung 5:	Statische Hindernisse von SAM 1.5. ....	25
Abbildung 6:	Dynamisches Hindernis von links nach rechts bewegend .....	26
Abbildung 7:	Schematische Zeichnung des Untersuchungsaufbaus.....	32
Abbildung 8:	Rating Scale Mental Effort <sub>ATEO</sub> und Rating Scale Effort <sub>ATEO</sub> .....	69
Abbildung 9:	Mittelwert (M) und Standardabweichung des Flächenfehlers beim Tracking in Millionen Pixel (Mio px) pro Durchgang (D) sowie über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0 .....	79
Abbildung 10:	Mittelwert (M) und Standardabweichung der benötigten Zeit zum Tracking in Sekunden (s) mit und ohne Ausreißer pro Durchgang (D) sowie über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0.....	81
Abbildung 11:	Mittelwert (M), Standardabweichung (SD) und Standardfehler des Mittelwerts (SE) der benötigten Zeit zum Tracking in Sekunden (s) mit und ohne Ausreißer für den kurzen (kTP) und langen (ITP) Tracking-Pfad .....	83
Abbildung 12:	Mittelwert (M) und Standardabweichung (SD) der subjektiven Beanspruchung von MWB 1 und 2 nach der RSE <sub>A</sub> für den kurzen (kTP) und langen (ITP) Tracking-Pfad sowie über beide hinweg .....	84
Abbildung 13:	Mittelwert (M) und Standardabweichung (SD) der subjektiven Komplexität nach MWB 1 und 2 für den kurzen (kTP) und langen (ITP) Tracking-Pfad sowie über beide hinweg.....	84
Abbildung 14:	Mittelwert (M) und Standardabweichung (SD) der subjektiven Instruktionsverständlichkeit von MWB 1 und 2.....	86
Abbildung 15:	Mittelwert (M) und Standardabweichung des Selbsturteils der MWB worauf beim Tracking mehr geachtet wurde je Messzeitpunkt und über alle Messzeitpunkte hinweg.....	88



Abbildung 16:	Mittelwerte (M) und Standardabweichungen (SD) der Beurteilungen der eigenen Präferenz und der des anderen MWBs hinsichtlich einer schnellen oder genauen Objektregelung .....	90
Abbildung 17:	Mittelwert (M) und Standardabweichung (SD) des Selbsturteils der Stimmung der beiden MWB von den Messzeitpunkten vor (MPrä2) und nach (MPost) den messrelevanten kooperativen Tracking-Durchgängen sowie über diese Messzeitpunkte hinweg .....	92
Abbildung 18:	Graphical User Interface der ATEO Automatiksoftware .....	109
Abbildung 19:	Schematische Zeichnung des ATEO LAB Systems (ALS) .....	115
Abbildung 20:	Funktionsklassen des Modells „Arten und Stufen von Automation“ nach Parasuraman et al. (2000) und das Ausmaß der Automation nach Onnasch, Wickens, Li und Manzey (2014) .....	128
Abbildung 21:	U-förmiger Verlauf der Gesamtkosten von Automation nach Kaber et al. (2009, S. 617, ins Deutsche übersetzt durch die Autorin) .....	132
Abbildung 22:	„Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme“ mit der Output-Variable „Teamleistung“ erweitert um Feedbackschleifen (Original s. Rocholl, 2010).....	155
Abbildung 23:	Schematische Zeichnung eines exemplarischen Untersuchungsaufbaus .....	193
Abbildung 24:	Chronologischer Ablauf der Untersuchung Leistungsdeterminante „Informationsverfügbarkeit“ .....	197
Abbildung 25:	Mittelwert (M) und Standardabweichung (SD) der, durch die Entwickler konzipierten, Anzahl an Automaten, Anzahl an Funktionen und durchschnittlichen Anzahl an Funktionen je Automat nach den Spezifikationsbogen sowie Anzahl an Funktionen nach der Bewertungsfunktionsmatrix pro Bedingung (B) der uV „Informationsverfügbarkeit“ .....	211
Abbildung 26:	Mittelwert (M) und Standardabweichung (SD) der Güte der Automatikkonzepte pro Bedingung (B) der uV „Informationsverfügbarkeit“ .....	212
Abbildung 27:	Mittelwert (M) und Standardabweichung (SD) des zweiten Residuums Güte der Automatikkonzepte (ohne gemeinsamen Varianzanteil der Variablen „Erfahrungsindex“, „Residuum Witzanzahl“ und „Leistungsmotiv HE“) pro Bedingung (B) der uV „Informationsverfügbarkeit“ .....	214
Abbildung 28:	„Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme inklusive Automationen“ – Darstellung der empirisch gefundenen Zusammenhänge zwischen den Input-, Prozessvariablen und der Output-Variable .....	220

## Abkürzungsverzeichnis

AAF	ATEO Automation Framework (Automatiksoftware)
AAFs	ATEO Automation Frameworks (Automatiksoftware)
AA-S	ATEO Automatensoftware
AMD	ATEO Master Display
AMS-R	Achievement Motive Scale-Revised (Leistungsmotivskala)
ATEO	ArbeitsTeilung Entwickler Operateur
aV	abhängige Variable
aVn	abhängige Variablen
BFI	Big Five Inventory
B 1	Bedingung 1 der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“: Systembeschreibung
B 2	Bedingung 2 der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“: Systembeschreibung + verschiedene Informationsquellen zu SAM
B 3	Bedingung 3 der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“: Systembeschreibung + verschiedene Informationsquellen zu SAM + verschiedene Informationsquellen zum Operateur
D	Durchgang
ell	Gabelung eckige, lange Pfadoption nach links ausgerichtet
elr	Gabelung eckige, lange Pfadoption nach rechts ausgerichtet
FESE-R	Fragebogen zur Erfassung von Strategie und Expertise in Experimenten, Revision
FM	Furcht vor Misserfolg
Gs	VP selektiert nach unspezifischer Strategie
Gu	VP <u>nicht</u> selektiert nach unspezifischer Strategie
GUI	Graphical User Interface
GUIs	Graphical User Interfaces
H	Häufigkeit
HAA (HTA)	Hierarchische Aufgabenanalyse (Hierarchical Task Analysis)
HE	Hoffnung auf Erfolg
IKD	Instrument zur Kodierung von Diskussionen
kTP	kurzer Tracking-Pfad

---

LFA	Logfile-Analyse-Tool
ITP	langer Tracking-Pfad
M	Mittelwert
MD	Messzeitpunkt Durchgang
Mio	Millionen
MMS	Mensch-Maschine-System
ms	Millisekunden
MW	Mikrowelt
MWB	Mikroweltbewohner
MWBn	Mikroweltbewohnern
MWBs	Mikroweltbewohners
NASA-TLX	NASA Task Load Index
OA	Operateursarbeitsplatz
OA-GUI	Operateursarbeitsplatz Graphics User Interface
OAs	Operateursarbeitsplatzes
OA-S	Operateursarbeitsplatzsoftware
PNG	Portable Network Graphics
px	Pixel
r	Korrelationskoeffizient nach Pearson/ Effektgröße
rl	Gabelung runde, lange Pfadoption nach links ausgerichtet
rlr	Gabelung runde, lange Pfadoption nach rechts ausgerichtet
RS(M)E <sub>A</sub>	Rating Scale (Mental) Effort <sub>ATEO</sub> (angepasste Version der RSME im ATEO-Projekt)
SAM	Socially Augmented Microworld/ Sozial Angereicherte Mikrowelt
SAM <sub>Java</sub>	Socially Augmented Microworld in Java (Java-Implementation)
SD	Standardabweichung (Streuung)
SW	Skalenwert
uV	unabhängige Variable
uVn	unabhängige Variablen
VI	Versuchsleiter
VP	Versuchsperson
VPn	Versuchspersonen
XLS	eXcel-Sheet
XML	Extensible Markup Language

---

25lr	Statisches Hindernis erst links, dann rechts positioniert mit 25 % Pfadabdeckung
25rl	Statisches Hindernis erst rechts, dann links positioniert mit 25 % Pfadabdeckung
50lr	Statisches Hindernis erst links, dann rechts positioniert mit 50 % Pfadabdeckung
50rl	Statisches Hindernis erst rechts, dann links positioniert mit 50 % Pfadabdeckung
$\tau$	Rangkorrelationskoeffizient Tau von Kendall

# 1. Einleitung

## 1.1 Einführung in das Forschungsthema und Ableitung der Ziele

Diese Dissertationsarbeit fand im Rahmen der zweiten Phase des Projekts ATEO (ArbeitsTeilung Entwickler-Operateur<sup>1</sup>) im Graduiertenkolleg prometei (Prospektive Gestaltung von Mensch-Technik-Interaktion) statt. Organisatorisch war prometei an das Zentrum Mensch-Maschine-Systeme (ZMMS) der Technischen Universität Berlin angegliedert und wurde von der Deutschen Forschungsgemeinschaft (DFG) gefördert. ATEO war in prometei organisatorisch eingebettet im Forschungsschwerpunkt 8: „Funktionsteilung Mensch-Maschine und ArbeitsTeilung Entwickler-Operateur – 2 Perspektiven auf Mensch-Maschine-Systeme“, welcher von Prof. Dr. Wandke, dem Lehrstuhlinhaber für Ingenieurpsychologie/ Kognitive Ergonomie der Humboldt-Universität zu Berlin, geleitet wurde.

Das Projekt ATEO fokussierte die Optimierung der Funktionsteilung (engl. Function Allocation) zwischen Operateur und Maschine in einem Mensch-Maschine-System (MMS). Bei der Funktionsallokation handelt es sich um die Zuweisung der Aufgaben an den Operateur oder die Maschine (Fuld, 1993). Im traditionellen Ansatz wird die Leistung eines Operators bei der Überwachung und Führung eines Systemprozesses (s. Supervisory Control) in Echtzeit analysiert und der Leistungsfähigkeit einer Automation gegenübergestellt, welche die gleiche Aufgabe ausübt (Parasuraman et al., 2000). Dieser Vergleich der Leistungsfähigkeit zwischen Operateur und Maschine ist ein alter (s. Fitts, 1951), aber in der Praxis wenig Nutzen bringender, Ansatz zur Erreichung einer adäquaten Funktionsteilung in einem MMS. Seit der Industrialisierung schreitet die technologische Entwicklung kontinuierlich und immer rasanter voran (Sheridan, 2000), wohingegen die Menschheit seit Jahrhunderten in ihrer kognitiven und physischen Leistungsfähigkeit keine erwähnenswerten Entwicklungssprünge zu verzeichnen hat (Bainbridge, 1998). Überspitzt ausgedrückt, ist somit ein heute aufgestellter, direkter Vergleich von Operateur und Maschine morgen bereits veraltet. Zur Gewinnung von langfristig gültigen Konzepten für die Funktionsteilung zwischen Mensch und Maschine ist daher ein gewinnbringenderer Ansatz von Nöten. Deshalb verfolgte das Projekt ATEO einen bislang von der Forschungsgemeinschaft lediglich theoretisch betrachteten (s. Hoc, 2001; Norman, 2007), jedoch potentiell nachhaltigeren Ansatz (Details s. Wandke & Nachtwei, 2008). Es handelt sich dabei um eine Perspektiverweiterung des Themas „Funktionsteilung zwischen Mensch und Maschine“ (Kain, 2008). Neben den Operateuren von komplexen, dynamischen, technischen

---

<sup>1</sup> Wenn vom „Operateur“ geschrieben wird, ist immer ein menschlicher Akteur gemeint.

Systemen wird eine zweite wichtige Gruppe von Menschen in die Betrachtungen mit einbezogen (Gross, 2008; Nachtwei, 2008a). Es handelt sich dabei um die Entwickler solcher Systeme oder deren Teilkomponenten, den Assistenzsystemen und Automaten. Denn hinter der Qualität eines technischen Systems, wie komplex es auch sein mag, steht die Leistung der Systementwickler (Krinner & Gross, 2005; Norman, 2007; Wiczorek & Zinn, 2008). Sie treffen prospektiv Entscheidungen, wie sich ein System in verschiedenen Situationen zu verhalten hat (Wandke & Wetzenstein, 2004). Ein technisches System agiert nur so gut, wie dessen Entwickler in der Lage waren, dessen Systemprozesse und potentiellen -störungen zu antizipieren (Kain, 2008). Diese Antizipationsfähigkeit ist entscheidend für die Gestaltung der Interaktion zwischen einem Operateur und einem technischen System. Die Interaktion wiederum beeinflusst den Erfolg eines MMS. Somit können Entwickler für das Verhalten von Maschinen verantwortlich gemacht werden (Sarter, Woods, & Billings, 1997). Entwickler können jedoch selbst von einem nicht erwarteten Systemverhalten überrascht werden (engl. Automation Surprises, s. Sarter et al., 1997). So bieten Mensch-Maschine-Schnittstellen<sup>2</sup> von MMS im Normalfall zumeist eine adäquate Aggregation und Visualisierung von relevanten Informationen zur Unterstützung der Operateure an. Jedoch in Ausnahmesituationen, etwa bei erheblichen Zeitdruck, erweist sich diese Informationsvermittlung oft als unzureichend (Cummings & Thornburg, 2011). Solche Ausnahmesituationen wurden von Entwicklern nicht antizipiert, weshalb ein dafür adäquates Systemverhalten fehlt. Aus diesem Grund stand im Mittelpunkt der Forschungsbemühungen des Projekts ATEO die Erkundung und Optimierung der asynchronen Arbeitsteilung zwischen Entwicklern und Operateuren zum besseren Verständnis und zur Optimierung der synchronen Funktionsteilung zwischen Mensch und Maschine (Gross, 2008; Nachtwei, 2008a). Dafür beschäftigte sich das Projekt mit einer längst in Vergessenheit geratenen Frage (s. Bainbridge, 1983): Wie gut können zukünftige Systemereignisse prospektiv durch Entwickler gemanagt werden im Vergleich zu Operateuren, welche auf das gleiche Ereignis zum Zeitpunkt des tatsächlichen Geschehens reagieren (Wandke & Nachtwei, 2008)?

Entwickler und Operateure von komplexen, dynamischen, technischen Systemen weisen grundlegend andere Rahmenbedingungen für ihre Tätigkeit auf. Beide bringen deren Leistung in ein und dasselbe System ein, jedoch zu unterschiedlichen Zeitpunkten (Gross, 2008). Entwickler haben im Vergleich zum Operateur viel Zeit für die Bearbeitung einer Über-

---

<sup>2</sup> Mensch-Maschine-Schnittstellen stellen das Bindeglied zwischen dem Menschen und dem technischen System inklusive der Assistenzsysteme in Form einer zentralen Kommunikations- und Interaktionseinheit dar (Maier, 2014). Diese umfasst alle Eingabe- und Ausgabeinstrumente eines Systems, welche der Interaktion zwischen dem Benutzer und dem System dienen (J. Schwarz, 2007).

wachungs- und Regelungsproblematik (Gross, 2008; Rocholl, 2010). Dagegen müssen Operateure in der Lage sein, schnell (Krinner & Gross, 2005), teilweise sogar in Sekundenbruchteilen (Gross, 2008), auf Systemereignisse zu reagieren, wie bspw. Störfälle. Aufgrund der unterschiedlichen zeitlichen und räumlichen Nähe zu solch einem Ereignis bauen beide Menschengruppen von diesem ein unterschiedliches Situationsbewusstsein (engl. Situation Awareness, s. Endsley, 1995b) auf. Operateure stecken direkt im Geschehen. Dadurch können sie bei einem transparenten Prozessverlauf ein hohes Situationsbewusstsein entwickeln. Welches es ihnen wiederum ermöglicht, ein Ereignis angemessen zu diagnostizieren und darauf zu reagieren respektive zeitnahe Ereignisse zu antizipieren (Doane, Sohn, & Jodlowski, 2004). Dagegen müssen Entwickler Prozessverläufe eines zukünftigen MMS samt unkritischer und kritischer Schwankungen schon lange im Voraus antizipieren, um ein adäquates System oder Teilkomponenten, wie Assistenzsysteme und Automatiken, entwickeln zu können. Durch den großen zeitlichen Abstand bauen sie i. d. R. ein geringeres Situationsbewusstsein auf, als es bei den Operateuren des Systems zu verzeichnen ist. Dieses Situationsbewusstsein ist von der Antizipationsgüte der Entwickler abhängig. Auf der Antizipation der verschiedensten Aspekte des zu unterstützenden Systemprozesses beruht die Entscheidung der Entwickler, welche Funktionen notwendig sind und welche davon automatisch, technisch unterstützt oder durch den Operateur ausgeführt werden sollen (Kain, 2008). Eine Systemstörung kann zu einer Behinderung, negativen Beeinflussung oder Unterbrechung des optimalen Prozessverlaufs führen (Badke-Schaub & Frankenberger, 2004). Je mehr und je detaillierter die normalen und kritischen Prozessverläufe von den Entwicklern vorhergesehen werden, desto höher kann die Angemessenheit des Systems oder dessen Teilkomponenten ausfallen. Die Entwickler spielen somit neben dem Operateur eine wichtige Rolle, wenn es um die Effizienz und Sicherheit eines MMS geht. Eine Tatsache, welche bereits Bainbridge (1983, 1987) in ihrer Arbeit „Ironies of Automation“ erwähnte.

Die Leistungsfähigkeit von Operateuren und Entwicklern wird durch bestimmte Faktoren beeinflusst, den sogenannten Leistungsdeterminanten (Kain, 2008). Hierbei kann es sich bspw. um deren Persönlichkeitsdimensionen handeln. Daher war ein Ziel des Projekts ATEO die Identifizierung einiger dieser Leistungsdeterminanten der Operateure, welche eine Wirkung auf deren Überwachung und Regelung eines Systemprozesses haben (s. Nachtwei, 2011a; Wandke & Nachtwei, 2008). Ebenso galt das Interesse den Entwicklern. Die Identifizierung ihrer Leistungsdeterminanten, welche einen Einfluss auf den Entwicklungsprozess und das -ergebnis haben, stellt deshalb einen Schwerpunkt dieser Dissertationsschrift dar.

Das Endziel des Projekts ATEO war die Zusammenführung der Ergebnisse beider Forschungsstränge. Dies wurde durch empirische Leistungsvergleiche zwischen Operateuren und Automaten bei der Überwachung und Führung eines technischen, komplexen, dynamischen Systems operationalisiert (Niestroj, 2014). Diese Automaten wurden von professionellen Entwicklern unter den im Projekt identifizierten, optimalen Ausprägungen ihrer Leistungsdeterminanten („Teamgröße“, „Erfahrung“, „fachliche Teamzusammensetzung“ und „Informationsverfügbarkeit“) im Rahmen der vorliegenden Dissertationsarbeit konzipiert (s. Abschnitt 5.2.3). Ebenso agierten die Operateure in dem Leistungsvergleich beeinflusst von den im Projekt identifizierten optimalen Ausprägungen ihrer Leistungsdeterminanten (Nachtwei, 2011a; Niestroj, 2014). Dabei wurden sie teilweise durch adaptierbare Automationen unterstützt (Niestroj, 2014). Zu klären war die Frage, welche Funktionen besser durch einen Operateur, welche besser durch eine Automatik und welche idealerweise durch eine kooperative Automation ausgeübt werden sollten.

Für diese Untersuchungen wurde dieselbe Versuchsumgebung verwendet. Diese wurde eigens dafür im Projekt entwickelt. Die Erweiterung der Versuchsumgebung zur Erreichung einer angemessenen wahrnehmbaren Komplexität für einen fairen Vergleich zwischen Entwicklern und Operateuren stellte einen weiteren Schwerpunkt dieser Dissertationsarbeit dar.

Die Forschungsarbeiten des gesamten Projekts können als gewinnbringend für die prospektive Gestaltung von Mensch-Technik-Interaktion angesehen werden. Die Erkenntnisse über die Leistungsdeterminanten von Entwicklern und Operateuren, wie auch jene über die Funktionsteilung von Mensch und Maschine, können zukünftige Systementwicklungsprojekte von Beginn an positiv beeinflussen. Dadurch können u. a. eine höhere Güte der Entwicklungstätigkeit und somit eine noch kompetenzförderlichere und zuverlässigere Funktionsteilung zwischen Mensch und Maschine erreicht werden, als es heutzutage der Fall ist (Gross, 2008). Zudem bleiben die leistungsbeeinflussenden Faktoren von Entwicklern langfristiger bedeutsam als Erkenntnisse über die Leistungsfähigkeit von technischen Systemen. Menschen werden sich in den nächsten Jahrhunderten kaum in ihren kognitiven und physischen Fähigkeiten ändern.

## **1.2 Aufbau des Projekts ATEO**

Die Tätigkeiten im Rahmen dieser Dissertationsarbeit stellten Zwischenetappen auf dem Weg zum Endziel – dem empirischen Leistungsvergleich zwischen Entwicklern, über deren Automaten, und Operateuren – des ATEO-Projekts dar. Daher muss die vorliegende Arbeit im Kontext des Gesamtprojekts betrachtet werden. Insgesamt war das Projekt in drei Phasen aufgegliedert. Jede Phase wurde von zwei Doktoranden bearbeitet, wobei einer den Entwickler und der andere den Operateur im Fokus seiner Forschungstätigkeiten hatte. Diese



Dissertationsarbeit ist der Entwicklerperspektive in der zweiten Phase des ATEO-Projekts zuzuordnen. Die Forschungstätigkeiten in den einzelnen Phasen werden nachfolgend kurz dargestellt, wobei auf die Inhalte dieser Arbeit etwas ausführlicher eingegangen wird.

Die erste Phase (2004-2007) diente zum einen der Entwicklung der in den ATEO-Untersuchungen verwendeten Versuchsumgebung sowie einem dazugehörigen Operateursarbeitsplatz (OA) für den empirischen Vergleich zwischen Entwicklern und Operateuren (Gross, 2006, 2008; Gross & Nachtwei, 2006, 2007).

Des Weiteren wurde bereits auf Seiten der Entwicklerperspektive versucht Leistungsdeterminanten von Entwicklern empirisch zu ermitteln. Dabei konnte in einem Extremgruppenansatz gezeigt werden, dass die potentiellen Leistungsdeterminanten „Teamgröße“, „fachliche Teamzusammensetzung“, „Erfahrung der Teams“ und „Art der Information über das System“ kombiniert zu Assistenzsystemen führen, welche mehr menschliche Handlungsphasen unterstützen. Jedoch konnte keine der überprüften Variablen eindeutig als Leistungsdeterminante identifiziert werden. (s. Krinner, 2009)

In der zweiten Projektphase (2007-2010) wurden die bisherigen Forschungsbemühungen fortgeführt. Zum einen musste die Versuchsumgebung erweitert werden, um einen fairen Vergleich zwischen Entwicklern und Operateuren zu ermöglichen. Zum Ende der ersten Projektphase wies die Versuchsumgebung eine unzureichende Komplexität auf, um einen solchen Vergleich gewährleisten zu können. Die Weiterentwicklung wurde von der Autorin gemeinsam mit ihrem Projektpartner Nachtwei betrieben. Zudem wurde für den experimentellen Leistungsvergleich zwischen Operateuren und Entwicklern eine Automatiksoftware mit Graphical User Interface zur Implementierung und Konfiguration von Automaten entwickelt. Ferner wurden Leistungsdeterminanten von Entwicklern und Operateuren identifiziert. Auf Seiten des Operateurs entstanden dadurch verschiedene Operateursarbeitsplätze, von denen der geeignetste für den Leistungsvergleich zwischen Entwicklern und Operateuren experimentell ermittelt wurde (s. Abschnitt 3.3.2 oder Nachtwei, 2011a). Auf Seiten der Entwickler und somit durch die Autorin wurden weitere potentielle Leistungsdeterminanten aus der Forschungsliteratur identifiziert, welche in einem „deskriptiven Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme“ mündeten (s. Rocholl, 2010). Des Weiteren wurde der Einfluss der potentiellen Leistungsdeterminante „Informationsverfügbarkeit“ auf den Entwicklungsprozess und das -ergebnis, den Automaten, systematisch experimentell überprüft. Im Rahmen dieser Untersuchung wurden die Wirkungen einiger weiterer potentieller Leistungsdeterminanten ermittelt, hierzu zählt bspw. das Persönlichkeitsmerkmal „Gewissenhaftigkeit“. Die bei der Untersuchung gewonnenen Automatenkonzepte wurden

hinsichtlich ihrer Güte bewertet, um das geeignetste für den Leistungsvergleich zwischen Entwicklern und Operateuren experimentell zu ermitteln. Dafür wurde von der Autorin ein „Bewertungsinstrument der Güte von Automatikkonzepten“ entwickelt.

In der dritten Projektphase (2010-2013) wurde die Erfahrung der Entwickler und die zur Verfügung stehende Zeit für die Entwicklung als potentielle Leistungsdeterminanten der Entwickler experimentell betrachtet. Dabei zeigte sich kein signifikanter Leistungsunterschied zwischen erfahrenen Entwicklern und Studenten, sofern Letzteren mehr Zeit zur Konzeption von Automaten zur Verfügung stand. (s. Niestroj, 2014)

Darüber hinaus beschäftigte sich diese Phase experimentell mit Ansätzen der kooperativen Automatisierung (s. Hakuli, et al., 2009). Dafür wurden adaptierbare Automaten in den Operateursarbeitsplatz implementiert. Aus diesen Untersuchungen wurden Empfehlungen für eine flexible Funktionsallokation in realen komplexen, dynamischen, technischen Systemen abgeleitet. (s. von Bernstorff, 2014)

Ebenso erfolgten die experimentellen Leistungsvergleiche zwischen Entwicklern, über deren Automaten, mit manuell tätigen Operateuren sowie Operateuren, welche durch kooperative Automationen unterstützt wurden. Hierbei zeigte sich, dass die Automaten der Entwickler, aber auch die Operateure, welche durch kooperative Automationen unterstützt wurden im Vergleich zu manuell tätigen Operateuren die Genauigkeit des Systemprozesses signifikant verbessern. (s. Niestroj, 2014)

### **1.3 Aufbau der Arbeit**

Diese Dissertationsschrift teilt sich in zwei Themenblöcke auf. Bevor auf diese im Detail eingegangen wird, erfolgt in Kapitel 1 die Einleitung in die Forschungsthematik inklusive Darstellung des Rahmenprojekts ATEO mit dessen Problem- und Zielstellung.

Der erste Themenblock umfasst die Weiterentwicklung der ATEO-Versuchsumgebung, um einen fairen Leistungsvergleich zwischen Entwicklern und Operateuren zu ermöglichen. Der zweite Themenblock beinhaltet die Identifizierung der Leistungsdeterminanten der Entwickler.

Der erste Themenblock beginnt, zum besseren Verständnis dieser Arbeit, mit der Beschreibung der Ausgangslage und den speziellen Anforderungen an eine Versuchsumgebung im ATEO-Projekt (Abschnitt 2.1). Daraufhin wird kurz dargestellt, wie in der ersten Projektphase versucht wurde diese Anforderungen mit einer eigens entwickelten Versuchsumgebung zu erfüllen sowie welche Mängel, trotz aller Bemühungen, am Ende vorlagen (Abschnitt 2.2). Im Kapitel 3 wird das Vorgehen zur Optimierung dieser Versuchsumgebung präsentiert. Hierbei werden im Abschnitt 3.1 die vorgenommenen Veränderungen beschrieben. Dem folgt

im Abschnitt 3.2 die Darstellung des umfangreichen experimentellen Manipulation-Checks der vorgenommenen Veränderungen. Dieser umfasst vier Voruntersuchungen von Komponentenvarianten zur Wirküberprüfung auf die Versuchsumgebung (Abschnitt 3.2.1 und 3.2.2). Ebenso gehört dazu eine Vergleichsuntersuchung der alten und der neuen Versuchsumgebung, welche im Abschnitt 3.2.3 präsentiert wird.

Im zweiten Themenblock erfolgt eine Darstellung des aktuellen Forschungsstands zum Thema „Automation“ und deren Entwickler (Kapitel 4). Hierbei werden die Chancen und Risiken von Automation sowie die Merkmale von guten Entwicklern und Entwicklungsprozessen vorgestellt. Im sich anschließenden Kapitel 5 geht es im Speziellen um die Leistungsdeterminanten dieser Entwickler. Es wird zunächst, zum besseren Verständnis dieser Arbeit, ein Überblick über die Forschungsaktivitäten in der ersten Projektphase zur Identifizierung der Leistungsdeterminanten gegeben (Abschnitt 5.1). Hiernach wird im Abschnitt 5.2.1 das eigens entwickelte „Modell der Leistungsdeterminanten von Entwicklern“ als Basis für die weiteren Forschungsaktivitäten vorgestellt. Im sich anschließenden Abschnitt 5.2.2 wird das von der Autorin entwickelte „Bewertungsinstrument der Güte von Automaten“ präsentiert. Dieses soll ermöglichen, die Güte von experimentell gewonnenen Automatenkonzepten zu bestimmen, um so Rückschlüsse auf die Wirkung von potentiellen Leistungsdeterminanten ziehen zu können. Im letzten größeren Abschnitt (5.2.3) wird eine Untersuchung der potentiellen Leistungsdeterminante „Informationsverfügbarkeit“ dargestellt.

Den Abschluss dieser Arbeit bildet das sechste Kapitel mit einem Gesamtfazit über das Forschungsvorgehen, die gewonnenen Erkenntnisse und Instrumente sowie einem Ausblick über sinnvolle weitere Forschungsschritte.

**PART I**

**Weiterentwicklung**

**einer**

**Versuchsumgebung**

## **2. ATEO-Projektphase 1: Entwicklung einer Versuchsumgebung**

### **2.1 Ausgangslage**

Für die experimentellen Untersuchungen im ATEO-Projekt wurde eine Versuchsumgebung benötigt, welche eine ausreichende Tätigkeitsherausforderung für Entwickler und Operateure gleichermaßen bietet und darüber hinaus den Vergleich ihrer Leistungen ermöglichte (Gross, 2008). Ebenso musste die Versuchsumgebung die Merkmale eines komplexen, dynamischen, technischen Systems aufweisen (Wandke & Nachtwei, 2008), um eine hohe ökologische Validität zu besitzen. Dabei sollte diese nicht domänenspezifisch sein. Dadurch können die Ergebnisse des Projekts auf eine Vielzahl realer komplexer, dynamischer, technischer Systeme übertragen werden (Nachtwei, 2008a). Außerdem musste die Versuchsumgebung für Laboruntersuchungen geeignet sein und dabei die Ausschaltung oder Kontrolle einer Vielzahl an Störvariablen ermöglichen. Dadurch sollte eine experimentelle Kontrollierbarkeit (Nachtwei, 2008a) und somit eine ausreichende interne Validität gewährleistet werden (Nachtwei & Kain, 2010). Die Verwendung eines realen komplexen, dynamischen, technischen Systems kam daher nicht in Frage.

#### **2.1.1 Klassische Mikrowelt**

Technische Systeme werden immer umfangreicher und beinhalten teilweise mehrere hunderttausend miteinander interagierende Prozessvariablen (Burmester, Komischke, & Wüst, 2000). Die Auswirkungen dieser umfangreichen Variableninteraktionen sind für Entwickler immer schwerer bzw. überhaupt nicht mehr nachzuvollziehen (Sheridan & Parasuraman, 2006). Zudem weisen diese Systeme oftmals mehrere Ziele und damit einhergehend vielfältige, teilweise konkurrierende Aufgaben mit hohen Anforderungen für den Operateur auf, welche dieser unter hohem Zeitdruck ausführen muss (Endsley & Kaber, 1999; Kaber & Endsley, 1997, 1998, 2004). Diese Variablen- und Aufgabenvielfalt inklusive ihrer Interaktion ist experimentell nicht zu kontrollieren. Trotz systematischer Datenerhebung und -analysen würden keine sicheren kausalen Aussagen getroffen werden können, da die Ergebnisse durch eine Vielzahl weiterer Variablen (mit-)verursacht sein könnten (Bortz & Döring, 2006). Einige Wissenschaftler beschreiben dies als ein für die Forschungswelt typisches Dilemma (Brehmer, 1992; DiFonzo, Hantula, & Bordia, 1998). In der Feldforschung herrscht meist eine zu hohe Komplexität für eindeutige Ergebnisse, wohingegen durch die eher zu geringe Komplexität im Labor selten praxisrelevante Wirkungszusammenhänge aufgedeckt werden können (Brehmer, 1992). Für einige Forscher stellt die klassische Mikrowelt die Lösung dieses Konflikts dar (Brehmer, 2005; Brehmer & Dörner, 1993). Mikrowelten sind computerbasierte Simulationen von realen Sys-

temen, durch welche gleichzeitig eine hohe ökologische Validität und eine adäquate experimentelle Kontrolle von Störvariablen erreicht werden kann (Badke-Schaub, Hofinger, & Lauche, 2008; DiFonzo et al., 1998; Dörner, 1989; Köster & Mehl, 2000; Sauer, Wastell, & Hockey, 2000). Mittels Simulationen kann eine Vielzahl an realen technischen und physikalischen Prozessen nachgebildet werden (Brehmer & Dörner 1993). Deshalb wurde zu Beginn des Projekts entschieden, als Versuchsumgebung eine Mikrowelt in Form der Simulation von realen Überwachungs- und Führungsprozessen zu verwenden. Es existieren viele derartige Simulationen (s. Sauer, Wastell, Hockey, Crawshaw, & Downing, 2003; Sheridan, 2002). Sie sind in Anlehnung an reale Systeme entstanden und stellen ein eigenes in sich abgeschlossenes, funktionierendes System dar. Lediglich die untersuchungsrelevanten Aspekte der Realität werden hierbei abgebildet (Gonzalez, Vanyukov, & Martin, 2005; Waern & Canas, 2003). Dadurch werden die Eigenschaften des realen Systems so vereinfacht, dass relevante Zusammenhänge sichtbar und verständlich werden (Gross, 2008). Außerdem kann so die notwendige Komplexität und Dynamik der Realität abgebildet werden (DiFonzo et al., 1998; Frieling & Sonntag, 1999). Für eine hohe ökologische Validität einer Simulation müssen deren Eigenschaften so nah wie möglich an der Realität sein (s. Simulation Fidelity bei Hays & Singer, 1989). Die Realität stellen die komplexen, dynamischen, technischen Systeme der Überwachungs- und Führungsdomäne dar (bspw. Trinkwasseraufbereitungsanlagen, Fertigungstechnik, Kernkraftwerke, Müllverbrennungsanlagen, Schifffahrt, Luftfahrt etc.).

#### **2.1.1.1 Determinismus vs. Stochastik**

Zu Beginn des Projekts existierte keine klassische Mikrowelt, welche einen komplexen, dynamischen, technischen Prozess realitätsrelevant simulierte und die zuvor definierte Anforderung erfüllte, sowohl die Leistung von Entwicklern als auch Operateuren unter kontrollierten Laborbedingungen fair beleuchten zu können. Entweder sind diese Mikrowelten deterministisch oder stochastisch (Brehmer & Dörner, 1993). Deterministische Mikrowelten stellen einen geeigneten Ansatz zum Testen der Leistungsfähigkeit von Operateuren bei der Überwachung und Führung von komplexen, dynamischen, technischen Systemen dar. Das konnte bereits durch die Ergebnisse zahlreicher Untersuchungen zur Gestaltung von Mensch-Maschine-Schnittstellen und deren Auswirkungen auf die Leistung, das Situationsbewusstsein sowie die mentale Beanspruchung (engl. Mental Workload) der Benutzer bewiesen werden (Nachtwei & Meyer, 2011). Dagegen sind deterministische Mikrowelten für die Analyse der Leistungsfähigkeit von Entwicklern ungeeignet. Ein Entwickler kann bei einem solchen System, auch wenn dessen Prozesse noch so komplex und interdependent sind (Gross & Nachtwei, 2007), alle möglichen Situationen und die damit verbundenen Systemzustände ohne

Mühen antizipieren, wenn ihm hierfür genügend Zeit zur Verfügung steht. Dadurch wäre es ihm möglich, die perfekte Automation zu entwickeln. Die bis dato gut etablierten Simulationen von Überwachungs- und Führungsprozessen sind deterministisch, beispielhaft sei hier CAMS genannt (s. Sauer et al., 2000). Deren bisherige Anwendungsgebiete erforderten keine nicht-deterministischen Ereignisse, so dass dieser Aspekt der Realität nicht berücksichtigt wurde, gleichwohl diese Ereignisse in den realen Systemen zumeist vorkommen und Probleme verursachen können (s. Normal Accidents Theory von Perrow, 1999).

In einer stochastischen Mikrowelt führen zufällig generierte Ereignisse (Brehmer & Dörner, 1993) dazu, dass die Entwickler keine Chance haben, den Prozess vollständig zu antizipieren (Gross, 2008). Dadurch können sie nicht vorhersehen, welche Automationen zur Unterstützung eines Systems notwendig sind. Entwickler haben deshalb bei einer solchen Mikrowelt einen klaren Nachteil gegenüber dem Operateur, der jegliches Ereignis, selbst wenn es zufällig auftritt, in Echtzeit wahrnehmen und darauf reagieren kann (Krinmer, 2009).

In gewissem Sinne können Mikrowelten als Testitems angesehen werden. Jedes Item besitzt einen bestimmten Schwierigkeitsgrad (inverse Wahrscheinlichkeit das Item zu lösen, Bortz & Döring, 2006). Wenn ein Item zu einfach oder zu schwer zu lösen ist, fällt die Varianz der Ergebnisse gering aus. Jede Variation einer zu untersuchenden Leistungsdeterminante hätte dann keinen erfassbaren Einfluss. Eine Varianz in den abhängigen Variablen, wie der „Güte der Automatikkonzepte“ von Entwicklern, wäre somit kaum bis gar nicht vorhanden (Nachtwei & Kain, 2010). Daher muss für einen fairen Leistungsvergleich von Entwicklern und Operateuren die Schwierigkeit der als Versuchsumgebung zu verwendenden Mikrowelt in etwa gleich und zudem mit einer mittleren Ausprägung ausgelegt werden. Dafür wird ein gewisser Grad an Unvorhersehbarkeit des Systemprozesses benötigt, welcher im Nachhinein gut erklärt werden kann. Auch andere Forscher, etwa Salas (2008), sprechen ebenfalls von dem Bedarf nach einem neuen methodischen Ansatz für Feld- und komplexe Versuchssettings in der Forschung. Eine ausführlichere Diskussion dieser Thematik liefern Wandke und Nachtwei (2008).

## **2.2 Entwicklung einer Mikrowelt als ATEO-Versuchsumgebung**

Um die zuvor genannten Anforderungen an die Versuchsumgebung zu erfüllen, wurde eigens im Projekt eine geeignete Mikrowelt entwickelt. Das Vorgehen wird nachfolgend beschrieben. Dabei wird kurz auf die Programmierumgebung Squeak eingegangen. Dem folgt eine Darstellung der Entwicklung der Versuchsumgebung in der ersten Projektphase. Abschließend wird die empirisch unterstützte Fertigentwicklung im Rahmen dieser Dissertation präsentiert.

### 2.2.1 Squeak als Programmierumgebung

Die Mikrowelt sollte innerhalb von kurzer Zeit von programmierunerfahrenen ATEO-Projektmitgliedern entwickelt werden. Deshalb wurde diese in der Open Source Programmier- und Ausführungsumgebung Squeak umgesetzt (Guzdial, 2001; Ingalls, Kaehler, Maloney, Wallace, & Kay, 1997; Korienek, Wrensch, & Dechow, 2002). Squeak basiert auf der objekt-orientierten Programmiersprache Smalltalk. Durch das Morphic-Framework von Squeak sollten ohne Programmierkenntnisse auf einfache Art interaktive grafische Benutzeroberflächen geschaffen und angepasst werden können (Maloney, 2000). Allerdings offerierten die Einstellmöglichkeiten eines Morphs nur einen eingeschränkten Spielraum, so dass für die Ziele des Projekts letztlich doch Codearbeit in Smalltalk durch erfahrene Programmierer unumgänglich war (Burandt, 2007). Von erfahrenen Programmierern kann Smalltalk aufgrund des geringen Sprachumfangs schnell erlernt werden (Kosjar, 2012; Seid, 2012a).

Für die Versuchsumgebung wurden im Laufe der zweiten ATEO-Projektphase die Squeak-Versionen 3.7, 3.9 und 3.10 verwendet.

### 2.2.2 Versuchsumgebung SAM 1.0

Nachfolgend werden, zum Aufbau eines grundlegenden Verständnisses von der Mikrowelt, ihre wesentlichen Aspekte und ihr Entwicklungsstand in der ersten Projektphase beschrieben. Ausführliche Informationen zu den Überlegungen, Theorien und empirischen Befunden, auf denen die Entwicklung basiert, sind den Arbeiten von Gross (2006, 2008) und Nachtwei (2008a) und beiden zusammen (2006, 2007) zu entnehmen.

Dieser Mikrowelt liegt die Annahme zugrunde, dass sich das menschliche Verhalten dafür eignet, den geforderten Zufallsanteil zu verursachen. Deshalb wurden zwei Personen als fester Bestandteil des simulierten komplexen, dynamischen, technischen Systems in die Mikrowelt integriert (Gross, 2006). Aus diesem Sachverhalt leitet sich deren Name „Sozial Angereicherte Mikrowelt“ (kurz SAM, engl. Socially Augmented Microworld) ab (Gross, 2008). Die beiden Personen werden im Folgenden als Mikroweltbewohner (MWB) bezeichnet. Die menschliche Interaktion ist sehr komplex, vielschichtig und dynamisch (Wandke & Wetzstein, 2004), wodurch sich die Transparenz der Mikrowelt reduzieren und damit einhergehend die Komplexität der Mikrowelt erhöhen soll (Gross, 2008). Dabei geht es um die durch einen Entwickler oder Operateur wahrnehmbare Komplexität. Durch die Interaktionen der beiden MWB wird den Entwicklern, aber auch den Operateuren die Antizipation der Prozessverläufe der Mikrowelt erschwert (Nachtwei & Kain, 2010). Bereits Johansson, Persson, Granlund und Mattsson (2003) beobachteten, dass Personen im Umgang mit Mikrowelten



selbst für eine Vielzahl der entstehenden dynamischen Probleme verantwortlich sind. Nach Flemisch, Kelch, Löper, Schieben und Schindler (2008) ist es aufgrund der Verhaltenskomplexität des Menschen schwieriger, ein mentales Modell von diesem zu entwickeln als von deterministischen Prozessen. Auf eine Simulation der beiden Menschen wurde verzichtet, da dies hohe Analyse- und Implementierungsaufwände bedeutet hätte, ohne deren vielfältigen Interaktionen perfekt nachbilden zu können (s. Burkhard, et al., 2011).

Im Detail besteht SAM aus den beiden MWBn, welche kooperativ eine Tracking-Aufgabe durchführen. Dafür müssen diese gemeinsam ein rundes Objekt einen virtuellen Pfad vom Start bis zum Ziel entlang regeln (Gross, 2008). Beim kooperativen Tracking hat jeder MWB mittels eines Joysticks<sup>3</sup> 50 % Einfluss auf die horizontale und auf die vertikale Bewegung des Objekts (Wandke & Nachtwei, 2008). Das Objekt befindet sich auf der Vertikalen immer fix 65 Pixel über dem unteren Bildschirmrand und der darunterliegende Pfad wird mittels der Joystick-Eingaben bewegt (Gross, 2008). Dadurch ändert sich die Position des Objekts relativ zum Pfad. Lediglich die Objektposition auf der Horizontalen des Bildschirms kann tatsächlich durch die MWB beeinflusst werden (Gross, 2008). Bei jedem Programmschritt (alle 30<sup>4</sup> bis 49 ms vom Beginn bis zum Ende einer Tracking-Aufgabe) werden die aktuellen Auslenkungen der Joysticks<sup>5</sup> erfasst und in die Bewegung des Objekts respektive des Pfads umgerechnet (Bothe et al., 2009).<sup>6</sup>

SAM kann auf zwei Ebenen betrachtet werden. Eine Ebene ist die Innenansicht. Aus dieser Sicht können die MWB als Operateure gesehen werden, welche gemeinsam das Tracking-Objekt regeln. Die zweite Ebene ist die Außenansicht von SAM. Aus dieser Sicht handelt es sich bei den MWBn nicht um Operateure, sondern um Systemkomponenten, genauer um Regelungsinstanzen von SAM. Die Objektregelung durch die MWB stellt den Prozess von SAM dar. Der Prozessverlauf kann durch eine weitere Person, den tatsächlichen Operateur von

---

<sup>3</sup> Bereits Poulton (1974) empfahl für das Tracking die Nutzung eines Joysticks.

<sup>4</sup> Die 30 ms stellen das erwünschte Optimum für die Dauer eines Programmschritts dar, welche im Quellcode von SAM definiert sind. Aufgrund des hohen Aufwands bei der Umrechnung der Joystickausrückungen in die Objektbewegung kann dieses Optimum trotz sehr leistungsstarker Computer zumeist nicht umgesetzt werden. Um eine ausreichende Leistung von SAM für die Versuchszwecke im Projekt ATEO zu erreichen, darf ein Programmschritt jedoch maximal eine Zeit von 49 ms benötigen. Noch länger andauernde Programmschritte führen zu einer durch die MWB wahrnehmbaren Umsetzungsverzögerung der Joystickbewegungen sowie einem ruckligen Verschieben der einzelnen Spielbretter des Pfads. (Niestroj, 2009)

<sup>5</sup> Wertebereich der Joystickausrückung nach links oder hinten (zum Körper hin) = -1024 bis 0, nach rechts oder vorn (vom Körper weg) = 1024 bis 0 (Bothe, Hildebrandt, & Niestroj, 2009)

<sup>6</sup> Bei maximaler Auslenkung beider Joysticks nach vorn bewegt sich der Pfad unter dem Objekt vertikal mit der maximalen Geschwindigkeit von 20.48 Pixel pro Programmschritt (Hildebrandt, 2009; Niestroj, 2009). Im Gegensatz dazu verursachen die maximalen Auslenkungen beider Joysticks nach hinten einen Stillstand des Pfads unter dem Objekt. Die Ruhelage beider Joysticks auf der Vertikalen führt zu einer Bewegung des Pfads unter dem Objekt von 10.24 Pixel pro Programmschritt (Hildebrandt, 2009). Bei der horizontalen Ruhelage beider Joysticks dagegen bewegt sich das Objekt nicht auf der Horizontalen (Burandt, 2007). Die maximale Auslenkung beider Joysticks nach links oder rechts ergibt eine Objektbewegung von 10.24 Pixeln pro Programmschritt in die jeweilige Richtung.

SAM, oder Automaten allein oder gemeinsam<sup>7</sup> überwacht und geführt werden. Die Güte des Prozessverlaufs von SAM entspricht der gemeinsamen Leistung der beiden MWB bei der Objektregelung. Die Leistung wird über die Anzahl an Fehlern in Form von Pfad verlassen sowie der benötigten Zeit vom Start bis zum Ziel operationalisiert (Nachtwei, 2008a). Eine optimale Leistung beim Tracking erreichen die beiden MWB nur, wenn diese miteinander kooperieren und gemeinsam deren Handlungen koordinieren (Krinner & Polkehn, 2006). Ein Fehler oder eine zu langsame vertikale Objektregelung durch die MWB entsprechen einer Systemstörung (Wandke & Nachtwei, 2008). Diese muss durch einen Operateur direkt und von den Entwicklern indirekt über deren Automaten beseitigt werden. Zur Bewältigung der Systemstörung haben die Operateure und die Automaten die Möglichkeit, indirekt (bspw. den MWBn Hinweise, Warnungen, Feedbacks geben) oder direkt (bspw. die Objektkontrolle zu übernehmen oder die Aktionen der MWB zu blockieren), in den Systemprozess einzugreifen (Gross, 2008). Daher sind die Automaten aus Sicht der MWB Assistenzsysteme, welche deren Tracking-Aufgabe unterstützen.

Eine Tracking-Aufgabe kann als ein Regelkreis<sup>8</sup> verstanden werden (Gross, 2008; Poulton, 1974). Aufgrund der Dynamik der Tracking-Aufgabe müssen die MWB kontinuierlich die Ist-Objektposition an die Soll-Objektposition anpassen. Die Soll-Objektposition verändert sich wiederum permanent, da die Joystick-Eingaben der MWB zu einem Voranschreiten des Objekts auf dem Pfad führen. Der Verlauf des Pfads verändert sich kontinuierlich. Der Operateur greift direkt und die Entwickler indirekt über deren Automaten in den MWB-Objekt-Tracking-Umgebung-Regelkreis ein. Somit ist das supervidierte kooperative Tracking (Gross, 2008) dafür prädestiniert, eine Regelungstätigkeit abzubilden, wie diese in komplexen, dynamischen, technischen Systemen zu finden sind (Nachtwei, 2008a), ohne dabei ein spezifisches System darzustellen.

---

<sup>7</sup> Im Rahmen dieser Arbeit spielt eine Überwachungs- und Führungstätigkeit gemeinsam durch einen Operateur und Automationen noch keine Rolle. Diese Thematik wird erst in der dritten ATEO-Projektphase aufgegriffen (s. Niestroj, 2014; von Bernstorff, 2014).

<sup>8</sup> In der Norm IEC 60050-351: 2009-06 wird der Begriff „Regelung“ wie folgt definiert: „Vorgang, bei dem fortlaufend eine variable Größe, die Regelgröße, erfasst, mit einer anderen variablen Größe, der Führungsgröße, verglichen und im Sinne einer Angleichung an die Führungsgröße beeinflusst wird.“ (2009, S. 56). Dazu wird angemerkt: „Kennzeichen für das Regeln ist der geschlossene Wirkungsablauf, bei dem die Regelgröße im Wirkungsweg des Regelkreises fortlaufend sich selbst beeinflusst.“ (DIN IEC 60050-351, 2009, S. 56). Deshalb wird in dieser Arbeit nicht die umgangssprachlich wohl geläufigere aber falsche Beschreibung einer Trackingaufgabe „das Steuern eines Objekts“ verwendet. In der Norm IEC 60050-351: 2009-06 wird der Begriff „Steuerung“ wie folgt beschrieben: „Vorgang in einem System, bei dem eine oder mehrere variable Größen als Eingangsgrößen andere variable Größen als Ausgangsgrößen auf Grund der dem System eigenen Gesetzmäßigkeiten beeinflussen.“ (2009, S. 56). Dazu findet sich folgende Anmerkung: „Kennzeichen für das Steuern ist der offene Wirkungsweg oder ein geschlossener Wirkungsweg, bei dem die durch die Eingangsgrößen beeinflussten Ausgangsgrößen nicht fortlaufend und nicht wieder über dieselben Eingangsgrößen auf sich selbst wirken.“ (DIN IEC 60050-351, 2009, S. 57).

### **2.2.2.1 Komplexität von SAM 1.0**

Damit SAM als prototypische Simulation von realen komplexen, dynamischen, technischen Systemen bezeichnet werden kann, welche den Bedarf des Einsatzes von Operateuren und Automaten rechtfertigt, muss die Mikrowelt einen ausreichenden Grad an wahrnehmbarer Komplexität aufweisen. Dies sollte über die Kombination von Personen mit bestimmten Persönlichkeitseigenschaften als MWB-Paar sowie dem Aufbau des Tracking-Pfads erreicht werden. Nachfolgend werden diese beiden Vorgehensweisen näher beschrieben.

#### **2.2.2.1.1 Personenmerkmale zur Erreichung einer ausreichenden Komplexität von SAM**

Der Auswahl der Personen, welche als MWB-Paar fungieren sollen, liegt das im Projekt entwickelte Kooperation-Konflikt-Komplexitäts Modell (3K-Modell) zugrunde (Details s. Nachtwei, 2008a). Dieses postuliert, dass bestimmte Personenmerkmale einer Person beim Tracking Einfluss auf deren objektive Leistung nehmen. Werden zwei Personen als MWB-Paar eingesetzt, welche konträre Ausprägungen bei diesen Personenmerkmalen aufweisen, wird erwartet, dass dies zu einem objektiven Leistungsunterschied zwischen diesen führt. Ferner sollten die MWB zielorientiert und eigennützig handeln. Deshalb wird angenommen, dass diese beim kooperativen Tracking einen Leistungsvergleich vornehmen (s. Bowers & Benford, zitiert nach Nachtwei, 2008a). Durch den Leistungsvergleich sollten die MWB die Divergenz zwischen sich wahrnehmen, welche wiederum Konflikte zwischen ihnen verursachen sollte. Dabei wird vermutet, dass diese Konflikte zu nicht zielorientierten Verhaltensweisen der einzelnen MWB führen (s. Jehn, Chadwick & Thatcher, zitiert nach Nachtwei, 2008a). Außerdem wird angenommen, dass solche Verhaltensweisen nicht von den Entwicklern oder dem Operateur erwartet werden, wodurch diese den Prozessverlauf von SAM schlechter vorher-sagen können sollten. Dadurch würde für diese beiden Menschengruppen die wahrnehmbare Komplexität der Mikrowelt ansteigen. (Nachtwei, 2008a)

Ein nach dem 3K-Modell (Nachtwei, 2008a) geeignetes Personenmerkmal ist die Strategie. Daher wird diese zur Steigerung der wahrnehmbaren Komplexität von SAM verwendet. Die Strategie existiert in einer unspezifischen und spezifischen Form. Die unspezifische Strategie wirkt sich indirekt auf die Tracking-Leistung aus, durch die Tendenz einer Person sehr schnell oder sehr genau zu agieren. Hierbei handelt es sich um eine mittelfristig stabile Persönlichkeitsvariable. Das Kontinuum dieser Variable wird durch die beiden Pole Geschwindigkeits- und Genauigkeitsorientierung gebildet. Somit besteht ein trade-off-Verhältnis zwischen den beiden Leistungsmerkmalen Geschwindigkeit und Genauigkeit. Für die Ermittlung der unspezifischen Strategie einer Person wurde der „Fragebogen zur Erfassung

von Strategie und Expertise in Experimenten“ (FESE) von Nachtwei (2008a) entwickelt. (Nachtwei, 2008a)

In SAM wird die indirekte Wirkung der unspezifischen Strategie durch eine dazu passende, direkt wirkende, spezifische Strategie in Form einer Instruktion für die Tracking-Aufgabe (schnell vs. genau zu agieren) unterstützt. Aufgrund der konträren Ausprägungen der Personen eines MWB-Paares bei der unspezifischen und spezifischen Strategie sollten diese unterschiedliche Ziele aufweisen (Nachtwei, 2008a), welche sich wiederum in unterschiedlichen Verhaltensweisen widerspiegeln sollten. Ein MWB sollte dazu tendieren, das Tracking-Objekt eher schnell als genau den Pfad entlang zu führen, wohingegen der andere MWB die entgegengesetzte Tendenz aufweisen sollte. Um den sich laut dem 3K-Modell (Nachtwei, 2008a) daraus ergebenden Konflikt zwischen den MWBn noch zu schüren, wurde bei diesen die Annahme aufgebaut, jeweils dieselbe Instruktion erhalten zu haben.

Gross und Nachtwei (2006) konnten experimentell den Einfluss der Strategie (spezifische und unspezifische Form) auf die Leistung beim Solo-Tracking<sup>9</sup> bestätigen.

Durch dieses Vorgehen bei der Auswahl und Instruktion der MWB ist der gesamte SAM-Prozess scheinbar stochastisch und schwer zu antizipieren. Im Nachhinein jedoch können die durch Beobachtung, Logfiles und Fragebogen gewonnenen Daten unter Einbezug psychologischer Theorien, Methoden und Befunde analysiert, interpretiert und somit verstanden werden.

#### **2.2.2.1.2 Pfadelemente zur Erreichung einer ausreichenden Komplexität von SAM**

Eine angemessene wahrnehmbare Komplexität der Mikrowelt für den Einsatz in Untersuchungen mit Entwicklern und Operateuren im ATEO-Projekt sollte ebenso über die Form (bspw. den Krümmungsgrad) der einzelnen Pfadelemente und deren Zusammensetzung (etwa der Varianz in der Abfolge) erreicht werden. Je höher der sich dadurch ergebene Regelungsaufwand ausfällt, desto höher gestalten sich die Anforderungen an die Wahrnehmung und Reaktion der MWB. Das sollte wiederum zu einer geringeren Leistung (Genauigkeit und Geschwindigkeit) bei der Objektregelung führen. Gross (2008) hatte bereits experimentell geeignete Elemente für die Zusammensetzung der Tracking-Pfade von SAM identifiziert. Als Ergebnis bestehen diese Pfade aus Geraden sowie leichten, mittleren und starken Kurven, um eine hohe Varianz bei der Objektregelung zu erreichen und dadurch Lerneffekte zu erschweren. Zudem kommen in SAM an verschiedenen Pfadstellen Gabelungen mit zwei Pfadoptionen vor, welche eine Entscheidung für eine der beiden Optionen von den MWBn erfordern sollen (s.

---

<sup>9</sup> Beim Solo-Tracking führt ein MWB die Trackingaufgabe allein mit einem Joystick aus, mittels dem dieser 100 % Einfluss horizontal und vertikal auf die Objektregelung nimmt.

Abbildung 2). Bei zwei dieser Gabelungen ist eine Pfadoptio lang und einfach gestaltet, wodurch diese mit einer hohen Genauigkeit absolvierbar sein sollte, wohingegen die andere kurz und damit schnell, aber gleichzeitig schwerer bewältigt werden können soll. (Gross, 2008)

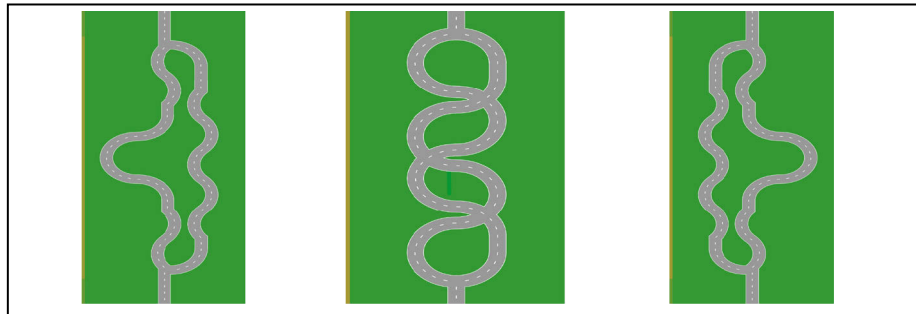


Abbildung 2: Die drei Gabelungsarten von SAM 1.0

Der erhöhte motorische Aufwand sowie Entscheidungen treffen zu müssen bei der Objektregelung, sollten wiederum zu Konflikte zwischen den MWBn verursachen. Dadurch sollte eine ausreichende, wahrnehmbare Komplexität der Mikrowelt für den Vergleich von Operateuren und Automaten erreicht werden.

#### 2.2.2.2 Mängel an SAM 1.0

Zum Ende der ersten Phase des ATEO-Projekts sollte SAM fertig entwickelt sein, um anschließend der Aufklärung der eigentlichen Fragestellungen des Projekts dienen zu können. Eine experimentelle Überprüfung von SAM zeigte jedoch, dass im Gegensatz zum Solo-Tracking die Tracking-Aufgabe von beiden MWBn gemeinsam zwar schneller, aber gleichzeitig mit weniger Fehlern absolviert wurde (Gross & Nachtwei, 2007). Die höhere Geschwindigkeit der Objektbewegung erfüllte die Erwartung an SAM für die Verwendung als komplexes, dynamisches, technisches System in den ATEO-Untersuchungen. Sie verursacht eine höhere Dynamik des Systemprozesses, wodurch auf Störungen im Prozess schneller durch einen Operateur oder die Automaten reagiert werden muss. Allerdings wird aufgrund der geringeren Fehlerrate als beim Solo-Tracking, der Einsatz von Operateuren oder Automaten zur Prozessregelung zu selten gefordert.

Da die Pfadverläufe in beiden Tracking-Bedingungen (solo vs. kooperativ) identisch waren, sah die Autorin die unterschiedlichen Tracking-Leistungen durch die MWB selbst bedingt. Möglicherweise hatten sich die gegensätzlichen Strategien nicht ausreichend im Tracking-Verhalten widerspiegelt. Dadurch entstand kein oder ein unzureichender Konflikt zwischen den MWBn. Zudem könnte es sein, dass beide MWB aufgrund von zu geringer Leistungsmotivation die Verantwortung für die kognitiv und motorisch aufwändigere genaue Objektregelung größtenteils bis ganz jeweils an den anderen MWB abgaben. Stattdessen wurde

ein schnelles Tracking ausgeführt. Ebenso könnte der Leistungsdruck durch den anderen MWB dazu geführt haben, die Objektregelung schneller auszuführen. Schnelleres Regeln bei gleicher Expertise müsste theoretisch zu einer höheren Fehlerrate führen. Allerdings könnten sich die Joystick-Bewegungen der beiden MWB so kompensiert haben, dass sich eine weniger fehlerhafte Objektführung ergab. Aufgrund der Befunde zu SAM musste die Autorin in ihrem Dissertationsvorhaben, anders als geplant, in einem ersten Schritt SAM optimieren. Ziel war es, zur Erhöhung der durch Entwickler und Operateure wahrnehmbaren Komplexität dieser Mikrowelt, mehr Störungen im SAM-Prozess zu erreichen.

### **3. ATEO-Projektphase 2: Weiterentwicklung der Versuchsumgebung**

Nachfolgend wird die empirische Weiterentwicklung von SAM 1.0 in der zweiten Projektphase ausführlicher beschrieben.

#### **3.1 Methodik der Weiterentwicklung von SAM**

Die Versuchsumgebung SAM 1.0 wurde inhaltlich und technisch weiterentwickelt, um mehr Störungen im Prozessverlauf und dadurch mehr Einsatzsituationen für einen Operateur oder Automaten zu schaffen. Dabei lag der Fokus auf der Erhöhung des Konflikts zwischen den MWBn, da das 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) einen Zusammenhang mit der Komplexität von SAM postuliert. Zusammen mit ihrem Projektpartner Nachtwei konzentrierte sich die Autorin auf vier Stellschrauben, die vielversprechend erschienen, Konflikte zwischen den MWBn steigern zu können. Hierbei handelt es sich um die Leistungsmotivation der MWB, die Glaubhaftigkeit der „Teamaufgabe“, die Entscheidungssituationen und der Joystick-Input der MWB bei der Objektregelung. So entstand die neue SAM-Version 1.5. In den nachfolgenden Abschnitten wird auf die vier Stellschrauben detaillierter eingegangen. Die Veränderungen an den letzten beiden Stellschrauben wurden von Niestroj (2008, 2009)<sup>10</sup> in Squeak implementiert. Die Autorin betreute diese Implementierungsarbeiten und führte die damit einhergehenden, zeitaufwändigen, manuellen Testungen hinsichtlich Funktionalität sowie Fehlerfreiheit von SAM durch.

Andere Möglichkeiten, wie Nebenaufgaben für die MWB, auf das Tracking-Objekt wirkende physikalische Kräfte bspw. in Form von Fliehkraft oder Seitenwind, unterschiedliche Pfaduntergründe, Höhenveränderungen oder variierende Objekteigenschaften wurden aus zeitökonomischen Gründen weiterhin außen vorgelassen. SAM 1.5 wurde jedoch für spätere Erweiterungen um diese und andere Aspekte flexibel programmiert, um weitere Möglichkeiten zur Komplexitätserhöhung schnellstmöglich implementieren zu können. Dadurch sollte ebenso die Nachnutzbarkeit dieser Versuchsumgebung für andere Forschergruppen selbst mit anderen Forschungsfragen gewährleistet werden.

##### **3.1.1 Steigerung der Leistungsmotivation der MWB**

Durch die Steigerung der Leistungsmotivation der MWB sollen diese animiert werden, ihre jeweilige Strategie (schnell vs. genau zu agieren) beim kooperativen Tracking durchzusetzen. Nachfolgend werden die dafür vorgenommenen Maßnahmen kurz dargestellt.

---

<sup>10</sup> Auf die mit dieser Dissertation verbundenen umfangreichen Implementierungsarbeiten wird in dieser Arbeit nicht eingegangen. Der interessierte Leser sei auf die Studentendarbeiten verwiesen.

### 3.1.1.1 Leistungsabhängige Entlohnung

Um die Leistungsmotivation der MWB zu steigern, wird erstmalig mit einer vermeintlich leistungsabhängigen Entlohnung für die Ausführung der Tracking-Aufgabe gearbeitet. Zu Beginn der Untersuchung wird den als MWB fungierenden Probanden mitgeteilt, dass die finanzielle Entlohnung für die Untersuchungsteilnahme erfolgsabhängig ist. Als Ausgangswert dienten zehn Euro. Für den Fall einer hohen Leistung wird den MWBn zwölf Euro und für den Fall einer niedrigen Leistung nur acht Euro als Entlohnung versprochen. Diese drei Stufen der leistungsabhängigen Entlohnung werden während der Untersuchung permanent auf einem A4 Blatt im Sichtfeld der MWB präsentiert (s. Abschnitt 3.2.1.2.3 und Anhang A.1.2). Dadurch soll die Wirksamkeit dieser Manipulation verstärkt werden. Tatsächlich erhalten alle MWB am Ende eines Untersuchungsdurchgangs zehn Euro. Zudem wird je Einsatz von SAM in einer Untersuchung ein Preis im Wert von 100 Euro für das MWB-Paar mit der besten Team-Performance ausgesetzt. Dies zusammen mit der leistungsabhängigen Entlohnung soll die extrinsische Motivation der MWB erhöhen (s. Weiner, 1994).

### 3.1.1.2 Leistungsfeedback

Zudem kann die Leistungsmotivation der MWB durch Feedback über deren eigene Leistung positiv beeinflusst werden (s. Weiner, 1994). In SAM 1.0 herrschte hinsichtlich des Feedbacks über die beiden Leistungskriterien (Genauigkeit und Geschwindigkeit) ein Ungleichgewicht. Den MWBn stand permanent eine Rückmeldung über die Genauigkeit des Objekt-Trackings zur Verfügung. Jedes Verlassen des Tracking-Pfads stellte einen Fehler dar. Für die Güte der Geschwindigkeit lag den MWBn jedoch nur die eigene Joystick-Auslenkung als Indiz vor. Um eine Balance zwischen der Transparenz hinsichtlich der Genauigkeit und der Geschwindigkeit des Tracking-Objekts zu erreichen, wurde in der neuen SAM-Version 1.5 ein visuelles Feedback über die Geschwindigkeit an die MWB umgesetzt. Dafür wurde in das Objekt eine 360°-Anzeige (Kreisdiagramm) der Objektgeschwindigkeit in Prozent implementiert (s. Abbildung 3). Dieser Anzeigeort wurde gewählt, damit die Geschwindigkeitsinformation sich genauso nah am Tracking-Objekt und damit im Hauptaufmerksamkeitsbereich der MWB befindet wie die Genauigkeitsinformation. Gleich relevante Informationen sollten nah beieinander (Vöhringer-Kuhnt, 2010) sowie im fovealen Sichtfeld positioniert werden (Johannsen, 1993), damit eine Reaktion auf ungewünschte Veränderungen angemessen schnell geschehen kann (Horrey, Wickens, & Consalus, 2006). Mit zunehmendem Abstand zum fovealen Sichtfeld sinkt die Wahrscheinlichkeit einer Blickzuwendung auf eine Information, da der perzeptive Aufwand zur Informationsaufnahme steigt (Sarodnick & Brau, 2006; Wickens,



2007). Zum anderen sollte dadurch die Gestaltung der Tracking-Umgebung so wenig wie möglich verändert werden, damit die Möglichkeiten zur Anzeige der visuellen Informationsermittlungen an die MWB durch den Operateur oder eine Automatik nicht eingeschränkt werden.

Im Stillstand ist die Objektfläche grün gefärbt. Die Geschwindigkeit wird durch eine schwarze Ausfüllung der Objektfläche in Analogie zu einem runden Tachometer vermittelt. Ein vollkommen schwarz ausgefülltes Objekt vermittelt maximale Geschwindigkeit. Leider konnte die Tachometeranalogie der Bewegung der Anzeige im Uhrzeigersinn nicht umgesetzt werden, da dafür keine Implementierungslösung gefunden werden konnte. Daher wurde eine geschwindigkeitsabhängige Ausfüllung des Objekts entgegen des Uhrzeigersinns umgesetzt.

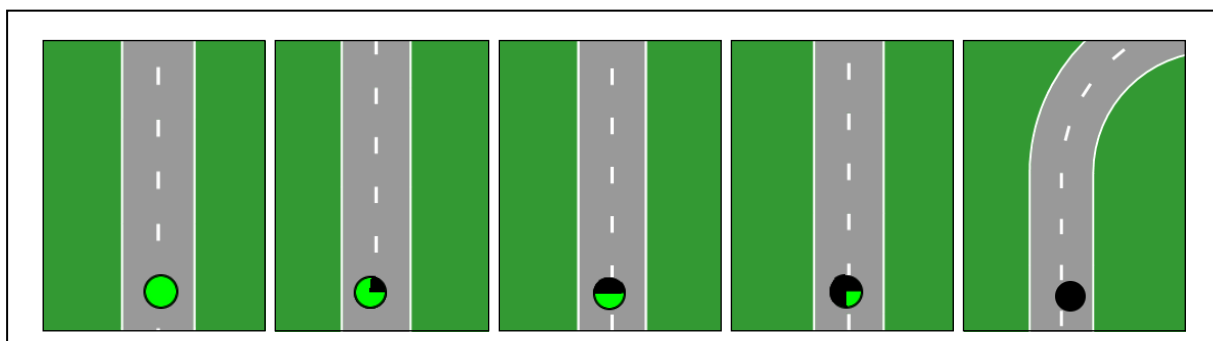


Abbildung 3: Geschwindigkeitsanzeige im Objekt in Form eines sich gegen den Uhrzeigersinn auffüllenden schwarzen Kreises in SAM 1.5, von links null Prozent (Stillstand) bis rechts 100 Prozent Geschwindigkeit

### 3.1.2 Erhöhung der Glaubhaftigkeit der „Teamaufgabe“

In SAM erhalten beide MWB als Aufgabe konträre spezifische Tracking-Strategien, wobei diesen vermittelt wird, dieselbe Aufgabe zu bekommen. Die Annahme ein gemeinsames Leistungsziel zu verfolgen, von dem jedoch die wahrnehmbare Objektregelung des jeweils anderen MWBs abweicht, sollte nach dem 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) den Konflikt zwischen den MWBn schüren.

Um die Chance in SAM 1.5 im Vergleich zu SAM 1.0 zu erhöhen, dass die MWB von einer gemeinsamen Teamaufgabe beim Tracking ausgehen, wird ein Trick angewendet. Dieser besteht darin, dass sich jedes MWB-Paar gemeinsam zum Beginn einer Untersuchung einen von fünf Briefumschlägen und damit seine Teamaufgabe auswählen darf. Dabei betont der Versuchsleiter (VI) mündlich, dass es sich um die Auswahl der Teamaufgabe handelt. Tatsächlich enthalten alle fünf Briefumschläge die gleiche Instruktion, so dass de facto keine Wahlmöglichkeit besteht.

Direkt im Anschluss an die Ziehung des Briefumschlages folgt für jeden MWB einzeln eine Test- und Trainingsphase. Dafür sitzt jeder MWB allein mit dem VI im Labor, während der andere MWB davor wartet. Für den ersten MWB bleibt der gewählte Briefumschlag die

gesamte Zeit vor ihm liegen. Nachdem der zweite MWB das Labor verlassen hat, erhält MWB 1 weitere Informationen für den Test- und Trainingsdurchgang, um diesen auf das kooperative Tracking vorzubereiten. Zudem dienen diese Informationen dazu, den MWB 1 von der Teamaufgabe abzulenken. Danach nimmt sich der MWB 1 den Briefumschlag und zieht seine vermeintliche Teamaufgabe zum Lesen heraus. Die Rolle des MWB 1 wird von Personen übernommen, welche bei der unspezifischen Strategie die Ausprägung geschwindigkeitsorientiert aufweisen. Daher befindet sich in dem Briefumschlag die schriftliche Instruktion, so schnell wie möglich zu steuern, welche durch eine grafische Visualisierung untermauert wird (s. Anhang A.1.1). Nach dem Lesen soll der MWB 1 die Instruktion für den anderen MWB wieder in den Briefumschlag stecken und diesen vor sich auf den Tisch legen. Dort bleibt dieser erneut für den MWB 1 permanent sichtbar liegen. Nach den Test- und Trainingsdurchgängen wird MWB 1 aufgefordert den MWB 2 hereinzuholen und selbst vor dem Labor zu warten. In dieser Zeit wird der Briefumschlag vom VI ausgetauscht. Als MWB 2 fungieren Personen, welche bei der unspezifischen Strategie die Ausprägung genauigkeitsorientiert besitzen. Deshalb erhält MWB 2 die Instruktion, so genau wie möglich zu steuern (s. Anhang A.1.1). Das weitere Vorgehen entspricht dem beim MWB 1.

Um zu vermeiden, dass die beiden MWB während der Untersuchung versehentlich diesen Trick aufdecken, wird sich eines zweiten Tricks bedient. Den MWBn wird verboten über ihre Teamaufgabe zu sprechen oder verräterische Mimiken und Gestiken zu vollführen, da die gesamte Untersuchung gefilmt und von unwissenden Beobachtern hinsichtlich des non-verbalen Verhaltens ausgewertet werden soll. Damit soll zudem die Kooperation zwischen den MWB erschwert werden. Foushee (1984, zitiert nach Hoc, 2001) zeigte den positiven Zusammenhang zwischen der Kommunikationsfrequenz von Personen und deren Leistung auf. Besonders in dynamischen Situationen ist verbale und non-verbale Kommunikation gleichbedeutend für die Kooperation (Hoc, 2001). Non-verbales Verhalten wird vor allem in kleinen Teams, wie Dyaden, zur Koordination von Aktivitäten, zur Erreichung der Ziele und damit zur Kooperation verwendet (Robertson & Zachary, 1990, zitiert nach Skjerve & Skraaning Jr., 2004). Damit der zweite Trick wirkt, befindet sich auf dem Monitor, auf dem die Tracking-Umgebung präsentiert wird, eine Webkamera. Diese signalisiert den MWBn während der Untersuchung über eine gut sichtbare, permanent rot leuchtende LED eine Videoaufzeichnung. Damit die MWB keinen Verdacht schöpfen, wird das Verbot von Sprache, Mimik und Gestik hinsichtlich der Tracking-Aufgabe diesen mit anderen Informationen zur Untersuchung gleichbedeutend intoniert vermittelt. Damit die MWB dieses Verbot einhalten, wird ihnen im Falle einer Missachtung zusätzlich der Verlust der finanziellen Entlohnung angedroht.

### 3.1.3 Mehr Entscheidungssituationen in Bezug auf Geschwindigkeit vs. Genauigkeit

Als eine weitere Stellschraube zur Erhöhung des Konflikts zwischen den MWBn wurde die Verwendung von Entscheidungssituationen für die MWB gesehen. Um diese bewältigen zu können, muss eine der beiden Tracking-Strategien der MWB umgesetzt werden. Im Folgenden werden die neuen Entscheidungssituationen in SAM 1.5 vorgestellt.

#### 3.1.3.1 Gabelungen

Entscheidungssituationen existieren bereits in SAM 1.0 in Form von Gabelungen (s. Abschnitt 2.2.2.1.2). Allerdings erscheinen diese ungeeignet, um von den MWBn eine Entscheidung zugunsten einer genauen oder schnellen Objektregelung treffen zu lassen. Zu dem Zeitpunkt an dem sich die beiden MWB spätestens für eine der beiden Pfadoptionen entscheiden müssen, sind diese noch nicht vollständig sichtbar. Die Entscheidung kann nur unter Unsicherheit, per Zufall oder nach dem Schema immer links respektive rechts getroffen werden. Somit erfolgt diese unabhängig von der Tracking-Strategie. Erst durch mehrfache Einsätze der Gabelungen während eines Untersuchungsdurchgangs können die MWB immer sichere Annahmen über den Verlauf der Pfadoptionen aufbauen und deren Tracking-Strategie in die Wahl einer Pfadoption einfließen lassen. Des Weiteren unterscheiden sich bei einer der verwendeten Gabelungen die beiden Optionen augenscheinlich nicht.

Das Konzept der Gabelung sollte in der neuen SAM-Version nicht aufgegeben werden, da es sich hierbei prinzipiell um ein gutes Element zur Erzwingung von Entscheidungen handelt. Um jedoch die Entscheidung für eine der beiden Pfadoptionen durch die Tracking-Strategie beeinflussen zu lassen, mussten die alten Gabelungen durch geeignetere Gabelungen ersetzt werden. Deshalb wurden neue Gabelungen entwickelt. Deren Länge wurde so angepasst, dass zum Zeitpunkt des Entscheidungserfordernisses beide Pfadoptionen in Gänze für die MWB zu sehen sind. Dadurch können die Eigenschaften der Pfadoptionen von der ersten Gabelung an in die Entscheidungsfindung mit einbezogen werden. Eine Pfadoption soll von dem geschwindigkeitsorientierten MWB 1 bevorzugt werden. Deshalb ist diese vorrangig geradlinig, dafür mit einer schmalen Pfadbreite gestaltet. Die andere Pfadoption soll von dem genauigkeitsorientierten MWB 2 präferiert werden. Dafür weist diese einen breiteren, jedoch längeren Pfad auf. Entscheiden sich die beiden MWB nicht für eine dieser beiden Pfadoptionen, sondern verfolgt jeder seine eigene Strategie, kommt es zu Fehlern.

Die Form des Pfads könnte einen Einfluss auf die Auswahl einer der Pfadoptionen haben. Daher wurden zwei Gabelungsvarianten gestaltet, eine mit rundem und eine mit eckigem Richtungswechsel des Pfads. Vor dem Monitor links sitzt immer der geschwindigkeitsorien-

tierte MWB 1 und rechts immer der genauigkeitsorientierte MWB 2. Dadurch könnte die Ausrichtung der Pfadoptionen einen Einfluss auf deren Auswahl haben. Deshalb wurden die zwei Gabelungen zudem gespiegelt, so dass insgesamt vier neue Gabelungsvarianten entstanden sind (s. Abbildung 4).

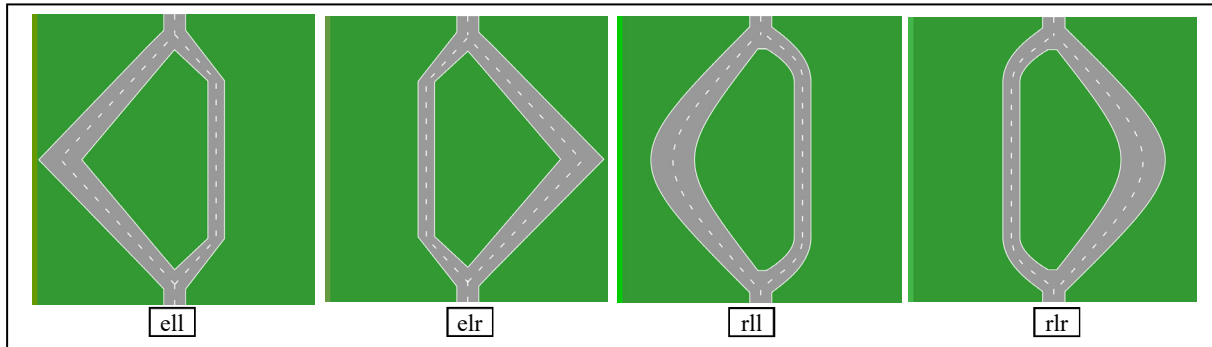


Abbildung 4: Gabelungen von SAM 1.5.

*Anmerkung.* Aus den beiden Eigenschaften Form und Ausrichtung der Gabelungen leitet sich deren Name ab. So steht der Name „ell“ für die Ausrichtung der eckigen, langen Pfadoption nach links.

### 3.1.3.2 Hindernisse

Neben den Gabelungen wurden neue Entscheidungssituationen in Form von statischen und dynamischen Hindernissen für die neue SAM-Version entwickelt. Eine Berührung des Objekts mit einem Hindernis stellt eine Kollision dar. Dadurch wird das Objekt links direkt neben den Pfad gesetzt, verbunden mit einem Stillstand für eine Sekunde. Eine Kollision verursacht somit einen durch die MWB wahrnehmbaren Nachteil für beide Leistungsdimensionen (Geschwindigkeit und Genauigkeit) gleichermaßen.

#### 3.1.3.2.1 Statische Hindernisse

Die statischen Hindernisse sind fest auf bestimmten Geraden der Tracking-Pfade positioniert. Diese decken von dem Pfad auf der linken oder rechten Seite einen prozentualen Teil (25 oder 50 %) ab. Diese Art des Hindernisses erfordert von den MWBn, das Tracking-Objekt um das Hindernis herum zu regeln. Die Art des Vorbeiführens sollte in Abhängigkeit der Tracking-Strategie variieren und kann somit zu Konflikten zwischen den MWBn führen. Die MWB können das Objekt am Hindernis schnell, mit der Gefahr einer hohen Abweichung vom Pfad, oder aber langsam und nahe daran vorbei regeln. Entscheiden sich die beiden MWB nicht für eine dieser beiden Strategien, sondern verfolgt jeder seine eigene Strategie, kann es sowohl zu Zeitverlusten als auch zu Fehlern kommen. Ein statisches Hindernis tritt immer mit einem zweiten statischen Hindernis auf der gegenüberliegenden Pfadseite in einem Abstand von ca. 250 Pixeln und mit der gleichen prozentualen Abdeckung des Pfads auf. Es handelt sich somit

eigentlich um einen Slalom. Durch die Kombination der vier verschiedenen statischen Hindernisse entstanden vier verschiedene Slaloms. Diese unterscheiden sich zum einen darin, wie viel Prozent diese von dem Tracking-Pfad versperren (25 vs. 50 %), zum anderen in der Reihenfolge der Positionierung der beiden Hindernisse (erst links, dann rechts vs. erst rechts, dann links; s. Abbildung 5).

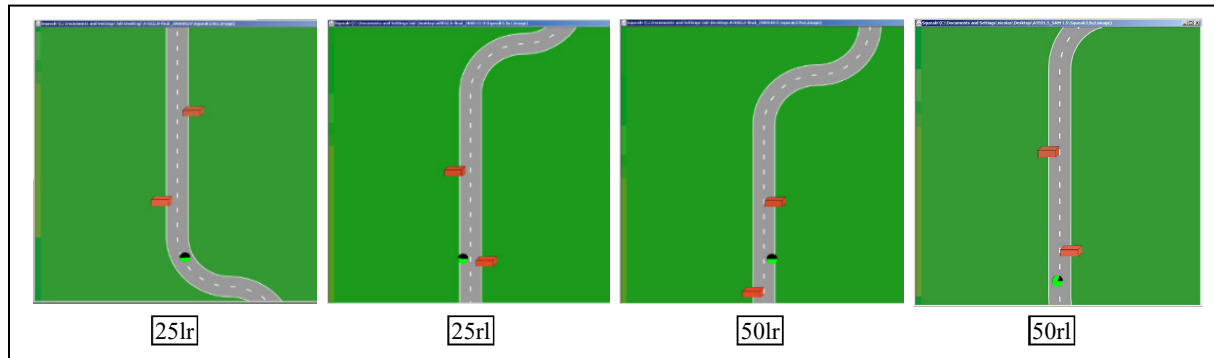


Abbildung 5: Statische Hindernisse von SAM 1.5.

*Anmerkung.* Der Name eines Hindernisses ergibt sich aus seinen Merkmalen. So vermittelt der Name „25lr“, dass beide statischen Hindernisse 25 % des Pfads abdecken und das erste Hindernis auf der linken, das zweite Hindernis auf der rechten Pfadseite platziert ist.

### 3.1.3.2.2 Dynamische Hindernisse

Für die Positionierung eines dynamischen Hindernisses in SAM ist lediglich eine Gerade<sup>11</sup> der Tracking-Pfade geeignet. Diese verläuft am rechten Bildschirmrand. Daher konnte in SAM 1.5 nur eine Variante eines dynamischen Hindernisses implementiert werden. Dieses Hindernis bewegt sich vom linken zum rechten Bildschirmrand, sobald es in den Sichtbereich der MWB gelangt (s. Abbildung 6). Bei einer Bewegung eines dynamischen Hindernisses von rechts nach links hätten den MWBn zu wenig Zeit zur Reaktion zur Verfügung gestanden. Die Geschwindigkeit des dynamischen Hindernisses wird adaptiv so an die Geschwindigkeit des Tracking-Objekts angepasst, dass es zu einer Kollision zwischen diesen beiden kommt, wenn die Objektgeschwindigkeit nicht durch die MWB verändert wird. In der Regel existieren für die MWB zwei Möglichkeiten, um eine solche Kollision zu vermeiden. Zum einen können sie die Objektgeschwindigkeit erhöhen, um zu versuchen, das Objekt vor dem Hindernis den Kollisionspunkt passieren zu lassen. Zum anderen können sie die Objektgeschwindigkeit so weit reduzieren, dass das dynamische Hindernis ohne Gefahr den Kollisionspunkt vor dem Objekt passiert. Somit stehen sich erneut die Geschwindigkeits- und Genauigkeits-Strategie konträr gegenüber. Um das Objekt nicht mit dem dynamischen Hindernis kollidieren zu lassen,

---

<sup>11</sup> Ein Tracking-Pfad setzt sich aus mehreren Spielbrettern zusammen, bei welchen es sich um GIF-Grafiken handelt.

müssen die beiden MWB gemeinsam nach einer der beiden Strategien handeln. In wenigen Situationen macht nur eine dieser beiden Reaktionen Sinn, wodurch eine Entscheidung zwischen den beiden Strategien nicht nötig ist. Das ist der Fall, wenn das Objekt bereits eine zu hohe oder zu niedrige Geschwindigkeit aufweist, so dass eine weitere Erhöhung oder Reduzierung dieser nicht eine Kollisionsvermeidung bewirken würde.

### 3.1.4 Manipulation des Joystick-Inputs

Das 3K-Modell postuliert, dass die Wahrnehmung eines objektiven Leistungsunterschieds zwischen den MWBn zu Konflikten führt (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1). Somit stellt die Schaffung oder Verstärkung eines Leistungsunterschieds zwischen den MWBn eine weitere Möglichkeit zur Steigerung des Konflikts zwischen ihnen dar. Eine unterschiedliche Leistung der MWB kann durch eine Manipulation der Joystick-Inputs der MWB erreicht werden. Deshalb wird an kurvigen Stellen des Tracking-Pfads der prozentuale Anteil des Joystick-Inputs von einem der MWB an der Objektregelung reduziert. Der andere MWB behält während dieser Zeit konstant 50 % Einfluss auf die Objektbewegung. Die Manipulation des Joystick-Inputs wird abwechselnd bei beiden MWBn durchgeführt.

Der prozentuale Anteil, um den der Joystick-Input eines MWBs an der Objektregelung reduziert wird, muss groß genug sein, um über die Objektbewegung von den MWBn wahrgenommen zu werden. Gleichzeitig darf dieser jedoch nicht so groß sein, dass er von den MWBn als Manipulation durch den VI oder als eine Softwarestörung wahrgenommen wird. Deshalb wurden mittels Blindtestungen durch die Autorin mit Hilfe von drei weiteren ATEO-Teammitgliedern für SAM 1.5 die Varianten 4 %, 8 %, 12 % und 16 % Reduzierung des 50%igen Joystick-Inputs eines MWBs ausgewählt.

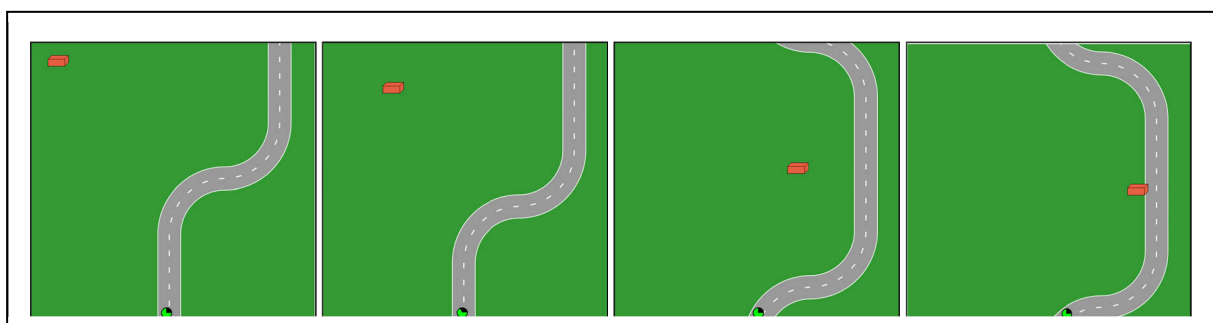


Abbildung 6: Dynamisches Hindernis von links nach rechts bewegend

## 3.2 Methodik der experimentellen Überprüfung der Weiterentwicklung

Durch Umsetzung und Implementierung aller zuvor genannten Möglichkeiten zur Steigerung des Konflikts zwischen den MWBn entstand SAM 1.5. Mittels dieser Version wurden

in drei kleinen Voruntersuchungen die geeignetsten zwei neuen Gabelungen, Hindernisse sowie Manipulationen des Joystick-Inputs zur Erhöhung des Konflikts zwischen den MWBn und damit der wahrnehmbaren Komplexität von SAM identifiziert. Die jeweils geeignetsten Varianten gingen in eine weitere SAM-Version 2.0 ein. Diese Version wurde mit SAM 1.0 experimentell verglichen, um zu überprüfen, wie erfolgreich die Weiterentwicklungen hinsichtlich der Steigerung der wahrnehmbaren Komplexität von SAM tatsächlich waren.

Nachfolgend wird dieser umfangreiche Manipulation-Check der Veränderungen an SAM beschrieben. Dafür werden zuerst die Voruntersuchungen und dann die Vergleichsstudie SAM 1.0 und 2.0 vorgestellt.

### **3.2.1 Voruntersuchungen: Neue SAM-Komponenten**

Aufgrund der vielen Gemeinsamkeiten der drei Voruntersuchungen zur Identifikation der geeignetsten zwei neuen Gabelungen, Hindernisse sowie Manipulationen des Joystick-Inputs und aus Platzgründen werden diese nachstehend zusammen beschrieben.

#### **3.2.1.1 Fragestellung**

In diesem Abschnitt wird die Herleitung der Forschungsfragestellung vorgestellt, die den drei Voruntersuchungen zugrunde lag. Hierbei handelt es sich um eine methodische Fragestellung zur Verbesserung des Untersuchungsinstruments „SAM“. Diese Fragestellung stellt keine inhaltliche Fragestellung zur Erreichung des ATEO-Forschungsziels dar.

##### **3.2.1.1.1 Herleitung der ersten Forschungsfragestellung**

Im Bemühen, den Konflikt zwischen den MWBn und damit die wahrnehmbare Komplexität von SAM zu erhöhen, wurden zum einen neue Pfadelemente in Form von Gabelungen und Hindernissen entwickelt, zum anderen wurde abwechselnd der Joystick-Input von einem der MWB an bestimmten Pfadstellen manipuliert (s. Abschnitt 3.1.3). Für die Gabelungen, die statischen Hindernisse und die Manipulation des Joystick-Inputs lagen jeweils vier Varianten vor. In die nächste Version von SAM sollten die jeweils zwei geeignetsten Varianten zur Steigerung des Konflikts zwischen den beiden MWBn integriert werden. Aus diesem Anliegen ergab sich folgende Fragestellung:

#### **Forschungsfragestellung 1:**

Welche zwei Gabelungen, statischen Hindernisse respektive Manipulationsstärken des Joystick-Inputs sind am geeignetsten zur Erhöhung der spezifisch auf Konflikte der MWB bezogenen Komplexität von SAM?

Für die Aufklärung dieser Fragestellung wurde ein exploratives Vorgehen gewählt. Es lagen keine Annahmen darüber vor, welche von den jeweils vier Varianten geeigneter sind, die wahrnehmbare Komplexität von SAM zu erhöhen.

### **3.2.1.2 Methoden**

#### **3.2.1.2.1 Operationalisierung der Variablen**

Der folgende Abschnitt dient dem Leser dazu, sich einen Überblick über die Operationalisierungen der unabhängigen und abhängigen Variablen zu verschaffen. In den Voruntersuchungen wurden zudem Kontrollvariablen erhoben. Diese wurden zudem in der Hauptstudie SAM 1.0 vs. SAM 2.0 durch die Autorin erfasst. Um den Leser nicht mit gleichen Analysen nur anhand unterschiedlicher Stichproben zu „langweilen“, wie auch aus Platzgründen, werden deren Analysen und die dazugehörigen Fragestellungen nur im Rahmen der Hauptuntersuchung anhand der größeren Stichprobe beschrieben. Dabei werden relevante Unterschiede zu den Analyseergebnissen der erfassten Daten der Kontrollvariablen in den Voruntersuchungen berichtet. Details zu diesen Kontrollvariablen der Voruntersuchungen (Erhebungszeitpunkt, Analyse, Ergebnisse) sind für den interessierten Leser im Anhang A.2 und 3 beschrieben.

##### **3.2.1.2.1.1 Operationalisierung der unabhängigen Variable**

In der ersten Voruntersuchung stellte die Gabelung die unabhängige Variable (uV) dar. Die uV der zweiten Voruntersuchung war das statische Hindernis. Die Manipulation des Joystick-Inputs, durch Reduzierung seines prozentualen Anteils an der Objektregelung beim jeweiligen MWB, diente in der dritten Voruntersuchung als uV. Diese drei uVn wurden jeweils systematisch variiert und vierfach gestuft (s. Abschnitt 3.1.3). Es handelte sich bei den drei uVn um einen Within-Faktor bzw. Messwiederholungsfaktor, d. h. einer Versuchsperson (VP) wurden beim Tracking alle vier Varianten der entsprechenden uV präsentiert. Nachfolgend sind die vier Stufen für jede uV dargestellt. Die in Klammern stehenden Abkürzungen der einzelnen Stufen der Faktoren werden im weiteren Verlauf der Arbeit, zur Erhaltung einer Übersichtlichkeit, anstelle der vollen Bezeichnungen verwendet.

##### **1. Voruntersuchung 1: uV „Gabelung“**

Die vier Stufen dieses Faktors in Form von vier Gabelungen unterschieden sich in der Form (Richtungswechsel des Pfads rund vs. eckig) und in der Ausrichtung (lange Pfadoption links vs. rechts):



*Gabelung 1: eckige, lange Pfadoption nach links ausgerichtet (ell)*

*Gabelung 2: eckige, lange Pfadoption nach rechts ausgerichtet (elr)*

*Gabelung 3: runde, lange Pfadoption nach links ausgerichtet (rll)*

*Gabelung 4: runde, lange Pfadoption nach rechts ausgerichtet (rlr)*

## **2. Voruntersuchung 2: uV „statisches Hindernis“**

Die vier Stufen dieses Faktors bestanden aus vier unterschiedlichen statischen Hindernissen. Bei diesen handelte es sich tatsächlich um einen Slalom, bestehend aus zwei statischen Hindernissen. Die statischen Hindernisse unterschieden sich durch den Anteil der Abdeckung des Pfads (25 oder 50 %) sowie der Position auf dem Pfad (erst links, dann rechts oder erst rechts, dann links):

*statisches Hindernis 1: erst links, dann rechts positioniert mit 25 % Pfadabdeckung (25lr)*

*statisches Hindernis 2: erst rechts, dann links positioniert mit 25 % Pfadabdeckung (25rl)*

*statisches Hindernis 3: erst links, dann rechts positioniert mit 50 % Pfadabdeckung (50lr)*

*statisches Hindernis 4: erst rechts, dann links positioniert mit 50 % Pfadabdeckung (50rl)*

## **3. Voruntersuchung 3: uV „Manipulation Joystick-Input“**

Die vier Stufen dieser uV entsprechen vier unterschiedlich starken Reduzierungen des prozentualen Anteils des Joystick-Inputs von einem der beiden MWB an der Objektregelung an kurvigen Stellen des Tracking-Pfads. Diese Manipulation des Joystick-Inputs betrifft abwechselnd einen der beiden MWB.

*Manipulation Joystick-Input 1: 4 % Reduzierung des Anteils an Objektregelung (4 %)*

*Manipulation Joystick-Input 2: 8 % Reduzierung des Anteils an Objektregelung (8 %)*

*Manipulation Joystick-Input 3: 12 % Reduzierung des Anteils an Objektregelung (12 %)*

*Manipulation Joystick-Input 4: 16 % Reduzierung des Anteils an Objektregelung (16 %)*

### **3.2.1.2.1.2 Operationalisierung der abhängigen Variablen**

#### **1. aV „Konflikt zwischen den MWBn“**

Der Konflikt zwischen den MWBn sollte sich in unterschiedlichem Objektregelungsverhalten der MWB widerspiegeln. Daher wurde die abhängige Variable (aV) „Konflikt zwischen den MWBn“ über die Differenzen der horizontalen und vertikalen Joystick-Auslenkung von MWB 1 und 2 und die Häufigkeiten der Wahl einer der beiden Pfadoptionen der Gabelungen operationalisiert. Je höher die Differenzen der Joystick-Auslenkungen und je ausgewogener die Wahl der beiden Pfadoptionen der Gabelungen sind, desto größer war der

Konflikt zwischen den MWBn.

*aV „Konflikt zwischen MWBn beim Objekt-Tracking“:*

- *Differenz der horizontalen Joystick-Auslenkung*
- *Differenz der vertikalen Joystick-Auslenkung*
- *Wahl der Pfadoptionen der Gabelungen*

## **2. aV „Genauigkeit“**

Ein unterschiedliches Objektregelungsverhalten der MWB sollte einen Einfluss auf deren gemeinsame Tracking-Leistung und damit auf die Güte des Prozessverlaufs von SAM haben. Die Leistung definiert sich über die Genauigkeit, mit der das Objekt den Tracking-Pfad entlang geregelt wird. Die Genauigkeit wurde durch die aV „NrSensor“ operationalisiert. Für die Ermittlung des NrSensors dienen in SAM 1.5 acht Sensoren, mit denen das Tracking-Objekt ausgestattet ist (s. Anhang G.1.1). Die Sensoren erfassen die Farbe, die sich unter ihnen befindet. Grau und weiß bedeuten, der Sensor liegt auf dem Pfad. Während grün dafür steht, dass sich der Sensor auf dem Randbereich befindet. Letzteres entspricht einem Fehler. Insofern kann ermittelt werden, wie viele der Sensoren nicht mehr auf dem Pfad liegen. Eine pixelgenaue Erfassung der Objektposition ist jedoch nicht möglich. Die aV „NrSensor“ spiegelt die Summe der Sensoren des Objekts wider, welche sich über alle Messungen vom Beginn bis zum Ende eines relevanten Pfadabschnitts nicht auf dem Pfad befunden haben. Es findet bei der Erfassung dieser Fehlerart keinerlei Umrechnung, sondern lediglich eine Aufsummierung, statt (Nachtwei, 2008a). Je weniger Fehler die MWB in einem Tracking-Durchgang verursachen, desto höher war deren Leistung.

*aV „Genauigkeit beim Objekt-Tracking“*

- *NrSensor*

## **3. aV „Geschwindigkeit“**

Die Leistung von Menschen spiegelt sich nicht nur in der Genauigkeit wider, sondern auch in der Geschwindigkeit, mit der eine Aufgabe ausgeführt wird (ISO/IEC 9126, 1991). Daher wurde zudem als aV die „benötigte Zeit“ zum Tracking vom Beginn bis zum Ende eines relevanten Pfadabschnitts operationalisiert. Die verwendete Einheit war Millisekunde. Je kürzer die benötigte Zeit für einen Tracking-Durchgang, desto höher war die Leistung der MWB.

*aV „Geschwindigkeit beim Objekt-Tracking“:*

- *benötigte Zeit*

Zur Erfassung wurden während eines Tracking-Durchgangs je Programmschritt (ca. alle 39 ms) die untersuchungsrelevanten Daten in ein automatisch angelegtes XML-Logfile geschrieben, wie bspw. die aktuelle Joystick-Auslenkung je MWB.

### **3.2.1.2.2 Erhebungsinstrument**

#### **3.2.1.2.2.1 Fragebogen FESE-R zur Vorselektion der VPn**

In den Voruntersuchungen wurde für die Vorselektion der VPn hinsichtlich ihrer unspezifischen Strategie (genauigkeits- vs. geschwindigkeitsorientiert) der „Fragebogen zur Erfassung von Strategie und Expertise in Experimenten, Revision“ (FESE-R) von Nachtwei (2008b) verwendet. Die Testgütekriterien der Kurzsкала besitzen adäquate Ausprägungen. Eine detailliertere Beschreibung des Messinstruments befindet sich im Anhang A.2.1.

#### **3.2.1.2.3 Aufbau der Voruntersuchungen**

Die Voruntersuchungen inklusive deren Probedurchläufe fanden an der Humboldt-Universität zu Berlin am Institut für Psychologie statt, in einem Raum des Lehrstuhls für Ingenieurpsychologie/ Kognitive Ergonomie. Dieser „reizarme“ Raum gewährleistete eine ungestörte Untersuchung. In ihm standen zwei Tische nebeneinander sowie drei Bürostühle.

Es wurde ein Drei-Computer-Aufbau verwendet (s. Abbildung 7). Für die Anwendung SAM 1.5 sowie der Logfile-Speicherung kam ein Computer (Betriebssystem: Microsoft Windows XP Professional SP2 64bit) zum Einsatz. Dieser befand sich rechts unter dem rechten Tisch, in ausreichender Entfernung zu den Sitzpositionen der MWB, so dass deren Beinfreiheit über die Dauer der Untersuchung nicht eingeschränkt wurde. An diesem Computer waren ein Monitor (19“, Bildschirmauflösung 1024 x 768 mit 75 Hz), eine Webkamera (Logitech Pro 9000 QuickCam mit 2-Megapixel und Autofokus), zwei Joysticks (Logitech Attack 3 Joystick), eine Maus sowie eine Tastatur angeschlossen. Der Monitor stand auf der linken Seite des rechten Tisches, vor dem die beiden MWB zum Ausüben der Tracking-Aufgabe Platz zu nehmen hatten. Der Monitor diente der Vorbereitung der Tracking-Aufgabe durch den VI und der Präsentation dieser für die MWB. Mittig auf dem Rahmen des Monitors befand sich die Webkamera. Als Eingabeinstrument für den VI dienten die Maus und die Tastatur. Die Tastatur befand sich während der Untersuchung nicht im Blick- und Griffbereich der MWB, so dass hierüber kein versehentlicher Abbruch des Ablaufs der Tracking-Aufgabe geschehen konnte. Den beiden MWBn stand die Maus zum Navigieren von einem zum nächsten Tracking-Durchgang und jeweils einer der beiden Joysticks zum Regeln des Tracking-Objekts zur Verfügung. Ein Joystick stand mittig vor der linken und der andere mittig vor der rechten Monitorhälfte. Dadurch sollten beide MWB einen guten Blick auf die Präsentation von SAM 1.5 haben,

ohne dass sie sich gegenseitig ihren Bewegungsfreiraum bei der Bedienung des Joysticks einschränkten. Zwischen den beiden Joysticks und dem Monitor befand sich permanent ein A4 Blatt, auf dem die leistungsabhängige Entlohnungsmodalität für die Teilnahme an der Untersuchung visualisiert war (s. Anhang A.1.2). Jeweils rechts auf den beiden Tischen stand ein Laptop (15“ Monitor, Betriebssystem: Microsoft Windows XP, Bildschirmauflösung 1280 x 800 mit 60 Hz) zur Darbietung der Fragebogen. Dadurch konnten die beiden MWB unabhängig und ungestört voneinander die Tracking-Durchgänge beurteilen. Auf den beiden Laptops war das FB-Tool des Lehrstuhls für Ingenieurpsychologie/ Kognitive Ergonomie der Humboldt-Universität zu Berlin zur Präsentation der Fragebogen und zur Answererfassung in einem Logfile installiert. An beiden Laptops war eine Maus für die MWB zum Navigieren durch die Fragebogen angeschlossen.

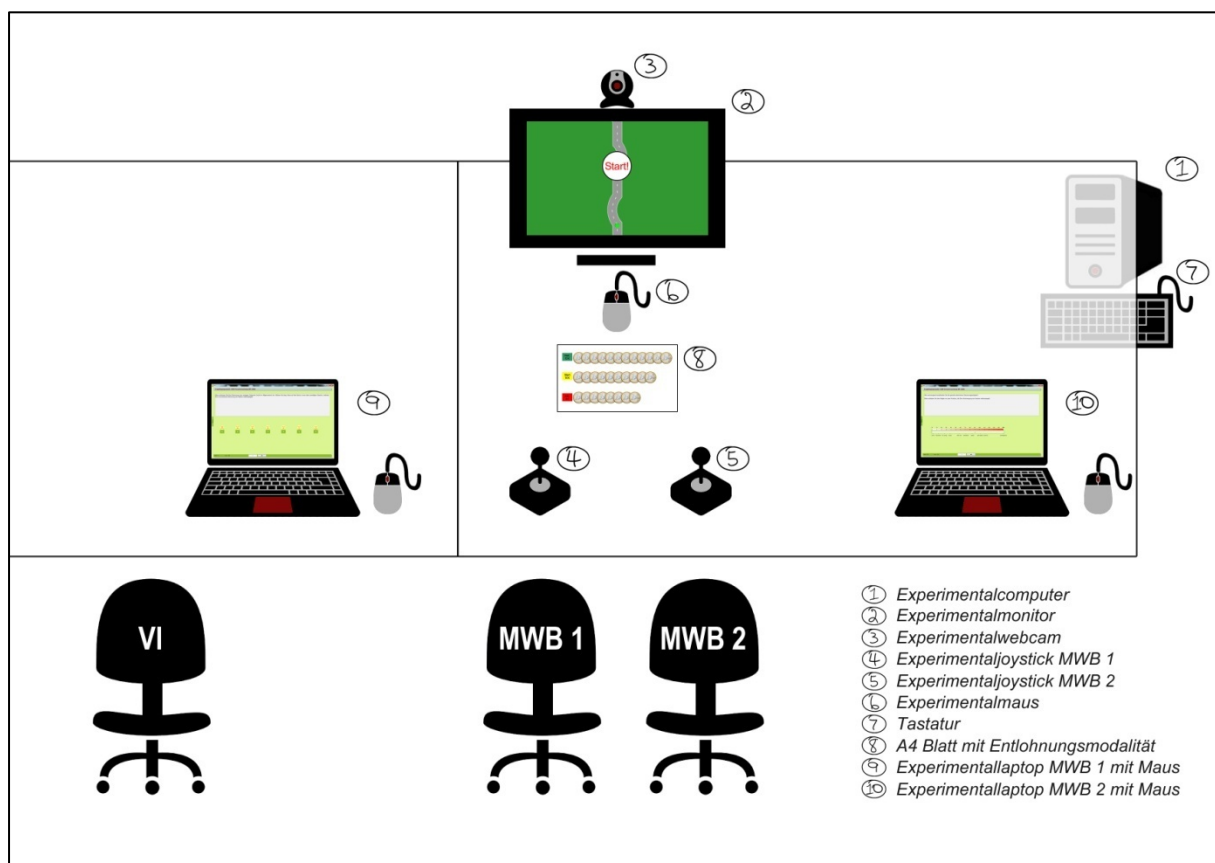


Abbildung 7: Schematische Zeichnung des Untersuchungsaufbaus

### 3.2.1.2.4 Experimentelles Design

Das experimentelle Design (s. Tabelle 1) der drei Voruntersuchungen war korrespondierend gestaltet. Lediglich die uV variierte. Insgesamt gab es sechs kooperative Tracking-Durchgänge, welche von je fünf Solo-Tracking-Durchgängen für jeden MWB flankiert wurden. Bei den

Tracking-Durchgängen Nummer 7, 8, 10, 12, 14, 16<sup>12</sup> handelte es sich um kooperative Tracking-Durchgänge. Die zu testenden vier Ausprägungen einer uV waren Bestandteil des Tracking-Pfads der Durchgänge 8, 10, 12 und 14. Bei der uV „Gabelung“ war in jedem der vier Durchgänge nur eine der jeweils vier Varianten dreimal implementiert. Bei der uV „statisches Hindernis“ kam in den vier Durchgängen je eine der statischen Hindernis-Varianten zweimal vor. Zudem erschien das dynamische Hindernis einmal in jedem der vier Durchgänge. Bei der uV „Manipulation des Joystick-Inputs“ kam je Tracking-Durchgang je eine der vier Varianten zweimal vor. Wobei je einmal MWB 1 und einmal MWB 2 von der Reduzierung des prozentualen Anteils an der Objektregelung betroffen war. Durch den mehrfachen Einsatz der verschiedenen Ausprägungen von jeder uV wurden wiederholte Messungen von deren Wirkung möglich.

Tabelle 1: Darstellung des jeweils aktiven MWBs in den 16 Tracking-Durchgängen eines Versuchsdurchlaufs

Nummer des Durchgangs	E	T	E	T	Messung											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
aktiver MWB	1	1	2	2	1	2	1 & 2	1 & 2	1	1 & 2	2	1 & 2	2	1 & 2	1	1 & 2

Anmerkungen. E = Durchgang zum Explorieren der Tracking-Aufgabe; T = Durchgang zum Trainieren der Tracking-Aufgabe

Um Reihenfolge- und Positionseffekte zu vermeiden, wurde ein experimentelles Design gewählt, welches durch die zyklische Permutation (s. Bortz & Döring, 2006) inspiriert war. Die vier Varianten einer uV wurden in vier verschiedenen Reihenfolgen eingesetzt (s. Tabelle 2, Tabelle 3, Tabelle 4). Eine vollständige Permutation war hinsichtlich eines ökonomischen Untersuchungsaufwands nicht möglich, da dies zu 24 verschiedenen Reihenfolgen je uV geführt hätte. Um zusätzlich Reihenfolge- und Positionseffekte durch den von der Manipulation des Joystick-Inputs betroffenen MWB zu vermeiden, war in jedem Tracking-Durchgang bei jeder der vier Varianten der Manipulation des Joystick-Inputs jeder MWB einmal zuerst betroffen. Zudem wurde diesbzgl. zwischen den Tracking-Durchgängen zwischen den MWBn gewechselt. Hieraus ergaben sich je Variantenreihenfolge zwei Reihenfolgen für den jeweils von der Manipulation des Joystick-Inputs zuerst betroffenen MWB (s. Tabelle 5).

<sup>12</sup> Durchgang 7 sollte ursprünglich der Identifizierung der Leistungs-Baseline der einzelnen MWB-Paare für das kooperative Tracking dienen. Mittels des Durchgangs 16 sollte ursprünglich ermittelt werden, welchen nachhaltigen Einfluss die Gabelungen, die Hindernisse respektive die Manipulationen des Joystick-Inputs auf das kooperative Tracking ohne das Vorhandensein dieser jeweiligen uVn haben. Ferner sollte damit überprüft werden, ob Zeiteffekte vorlagen, wie Lernen oder Konzentrationsveränderungen. Alle drei Aspekte sind miteinander konfundiert. Diese hätten aufgrund des Designs nicht einzeln überprüft werden können. Für die Analyse war ein sogenannter Prä-Post-Versuchsplan angedacht. Da nur sehr kleine Stichproben gewonnen werden konnten, wurde letztlich auf inferenzstatistische Analysen der objektiven Verhaltensdaten der MWB der einzelnen Voruntersuchungen verzichtet.

Tabelle 2: Variantenreihenfolgen der uV „Gabelung“

Reihenfolge	Durchgang 7	Durchgang 8	Durchgang 10	Durchgang 12	Durchgang 14	Durchgang 16
1	0	rlr	rlr	ell	elr	0
2	0	elr	rlr	rlr	ell	0
3	0	ell	elr	rlr	rlr	0
4	0	rlr	ell	elr	rlr	0

*Anmerkungen.* Bedingung Gabelung: rll = runde, lange Pfadoption nach links ausgerichtet; rlr = runde, lange Pfadoption nach rechts ausgerichtet; ell = eckige, lange Pfadoption nach links ausgerichtet; elr = eckige, lange Pfadoption nach rechts ausgerichtet; 0 = keine Manipulation

Tabelle 3: Variantenreihenfolgen der uV „statisches Hindernis“

Reihenfolge	Durchgang 7	Durchgang 8	Durchgang 10	Durchgang 12	Durchgang 14	Durchgang 16
1	0	25lr	25rl	50lr	50rl	0
2	0	50rl	25lr	25rl	50lr	0
3	0	50lr	50rl	25lr	25rl	0
4	0	25rl	50lr	50rl	25lr	0

*Anmerkungen.* Bedingung statisches Hindernis: 25lr = erst links, dann rechts 25 % Pfadabdeckung; 25rl = erst rechts, dann links 25 % Pfadabdeckung; 50lr = erst links, dann rechts 50 % Pfadabdeckung; 50rl = erst rechts, dann links 50 % Pfadabdeckung; 0 = keine Manipulation

Tabelle 4: Variantenreihenfolgen der uV „Manipulation des Joystick-Inputs“

Reihenfolge	Durchgang 7	Durchgang 8	Durchgang 10	Durchgang 12	Durchgang 14	Durchgang 16
1	0	4 %	8 %	12 %	16 %	0
2	0	16 %	4 %	8 %	12 %	0
3	0	12 %	16 %	4 %	8 %	0
4	0	8 %	12 %	16 %	4 %	0

*Anmerkungen.* Bedingung Manipulation Joystick-Input in %: 4, 8, 12, 16; 0 = keine Manipulation

Tabelle 5: Reihenfolgen der Manipulation des Joystick-Inputs der MWB

Reihenfolge	Durchgang 7	Durchgang 8	Durchgang 10	Durchgang 12	Durchgang 14	Durchgang 16
1	0	1, 2	2, 1	1, 2	2, 1	0
2	0	2, 1	1, 2	2, 1	1, 2	0

*Anmerkungen.* 1 = Manipulation des Joystick-Inputs von MWB 1; 2 = Manipulation des Joystick-Inputs von MWB 2; 0 = keine Manipulation

Das experimentelle Design stellte ein zweifaktorieller Versuchsplan mit Messwiederholung dar. Der messwiederholte Faktor, die jeweilige uV, und der Faktor „Variantenreihenfolgen der uV“ waren dabei jeweils vierfach gestuft. Somit handelte es sich um ein Mixed-Groups-Design (within und between). Bei der Voruntersuchung der uV „Manipulation des Joystick-Inputs“ kam der dritte messwiederholte Faktor „Reihenfolge betroffener MWB“ hinzu, welcher zweifach gestuft war.

### **3.2.1.2.5 Durchführung der Voruntersuchungen**

Die Voruntersuchungen fanden in Einzelsitzungen mit je einem MWB-Paar statt. Räumliche und technische Restriktionen waren hierfür ursächlich. Zudem sollte ein MWB-Paar die Untersuchung in einer ungestörten Atmosphäre ohne Zeitdruck absolvieren können. Die Rolle des Vls teilte sich die Autorin mit ihrem Projektpartner Nachtwei. Der Pfadaufbau, mit Ausnahme der neuen diagnostischen Elemente, der Ablauf und die Materialien waren der Untersuchung von SAM 1.0 (Gross & Nachtwei, 2007) entnommen und an die neuen Gegebenheiten angepasst worden. Zum einen hatten sich das Vorgehen und die Materialien bereits als zielführend und versuchsökonomisch erwiesen. Zum anderen konnte so die Vergleichbarkeit der SAM-Versionen dieser Projektphase mit SAM 1.0 gewährleistet werden.

#### **3.2.1.2.5.1 Probedurchläufe**

Jeder der drei Voruntersuchungen ging ein Probedurchlauf mit je einem VPn-Paar voraus, welche an den nachfolgenden Untersuchungen nicht mehr teilnahmen. Es zeigte sich, dass die Planungen der Voruntersuchungen geeignet waren für das Forschungsvorhaben. Es mussten keine Änderungen und somit weitere Probedurchläufe durchgeführt werden.

#### **3.2.1.2.5.2 Ablauf der Voruntersuchungen**

Der Untersuchungsablauf lässt sich in vier Phasen einteilen: Einführung der MWB, Test- und Trainingsphase, Hauptteil sowie Abschlussbefragung. Insgesamt müssen 16 Tracking-Durchgänge in SAM 1.5 absolviert werden. Bei den Durchgängen 1 bis 4 handelt es sich um kurze Pfade, für welche die MWB im Durchschnitt 30 Sekunden benötigten. Die Durchgänge 5 bis 16 beinhalten längere Pfade. Die MWB benötigten im Mittel drei Minuten, um diese zu absolvieren. Ein Untersuchungsdurchlauf dauerte im Durchschnitt 70 Minuten. Es folgt die Darstellung des chronologischen Ablaufs der drei Voruntersuchungen.

#### ***Phase 1: Einführung der MWB***

Die VPn kamen mit dem Wissen zur Untersuchung, an einer Tracking-Untersuchung teilzunehmen. Ein Untersuchungsdurchlauf startete mit der Begrüßung des VPn-Paares, welches die Rolle der MWB einnehmen sollte sowie einer kurzen Danksagung bzgl. der Teilnahmebereitschaft. Danach wurde ihnen kurz mündlich der Untersuchungszweck erläutert, nämlich die Wirkungstestung der Erweiterung von SAM. Zudem wurde den VPn mitgeteilt, dass jederzeit Verständnisfragen an den Vl gerichtet werden können. Ferner wurden sie gebeten, für die gesamte Untersuchung das Handy auszuschalten. Anschließend füllten die VPn eine Einverständniserklärung aus, in welcher die Anonymität der erhobenen Daten versichert wurde.

Danach erhielten die MWB eine kurze mündliche Instruktion vom VI sowie eine schriftliche Instruktion (s. Anhang A.2.4) am Monitor zum Untersuchungsablauf und zur Tracking-Aufgabe. Hiernach wurde ihnen die leistungsabhängige Entlohnungsmodalität (s. Abschnitt 3.1.1.1 und Anhang A.1.2) für die Teilnahme an der Untersuchung vorgestellt, wobei auf die Visualisierung auf dem Tisch verwiesen wurde. Zusätzlich wurde ihnen die Gewinnmöglichkeit von 100 Euro in Aussicht gestellt, wenn sie das MWB-Team mit der besten Tracking-Leistung sind (s. Abschnitt 3.1.1.1). Danach wurde darauf hingewiesen, dass ein Tracking-Durchgang nicht unterbrochen werden darf, es sei denn, einer der beiden VPn möchte die Teilnahme an der Untersuchung vorzeitig beenden. Dem folgte als Papierinstruktion die Manipulation der spezifischen Strategie (s. Abschnitt 3.1.2 und Anhang A.1.1). Dafür durften die MWB einen von fünf Briefumschlägen mit deren vermeintlichen „Teamaufgabe“ gemeinsam auswählen. Zur Erhöhung der Motivation der MWB diese Strategie beim Tracking anzuwenden, wurde ihnen mitgeteilt, dass die Wahrscheinlichkeit für die höchste Teilnahmevergütung und die Chance auf den Gewinn umso höher ist, je gewissenhafter die Strategie umgesetzt wird. Darüber hinaus wurden die beiden VPn darauf hingewiesen, nicht miteinander zu sprechen oder Gestiken und Mimiken zur Kommunikation über die „Teamaufgabe“ einzusetzen. Dies ging einher mit dem Verweis auf die Webkamera über dem Monitor und der Erläuterung des Zwecks der Videoaufnahmen, der Analyse des non-verbalen Verhaltens. Für den Fall einer Missachtung wurden die Streichung der Teilnahmeentlohnung und der Ausschluss an der Beteiligung des Zusatzgewinns angedroht.

### ***Phase 2: Test- und Trainingsdurchgang***

Es folgten für jeden MWB einzeln nacheinander, in Abwesenheit des anderen MWBs, zwei Solo-Tracking-Durchgänge. Der jeweils andere MWB wartete währenddessen vor dem Raum. Den MWBn wurde mitgeteilt, dass die beiden Durchgänge dem Kennenlernen des Trackings dienen und daher nicht mit in die Bewertung eingehen. Vor dem ersten Testdurchgang nahm jeder MWB die „Teamaufgabe“ zum Lesen aus dem Briefumschlag. Anschließend sollte diese wieder von den MWBn in den Briefumschlag gefügt werden. Begründet wurde dies damit, dass die „Teamaufgabe“ nicht permanent präsent auf dem Tisch liegen soll, um etwaige Ablenkungen zu vermeiden. Danach wiederholte der VI das Ziel der Untersuchung sowie die Regelungsmöglichkeiten des Objekts. Im Zuge dessen schlug der VI dem MWB vor, im ersten Durchgang die Objektregelung unabhängig von der „Teamaufgabe“ explorierend sowie mit Maximalauslenkungen auszuprobieren. Zudem erhielt der MWB eine kurze Startinstruktion am Monitor. Der MWB konnte durch Anklicken eines Startbuttons am Ende dieser Instruktion den ersten Solo-Tracking-Durchgang selbständig starten. Das freie Explorieren des MWBs wurde



vom VI durch paralleles Feedback mit Handlungsvorschlägen gefördert. Ein sich dem ersten Testdurchgang anschließender zweiter Solo-Tracking-Durchgang diente dem Training. Sowohl der VI wie auch die schriftliche Startinstruktion dieses Durchgangs wiesen den MWB darauf hin, diesmal die „Teamaufgabe“ beim Tracking zu berücksichtigen.

Während des Wechselns der beiden MWB wurde der Briefumschlag mit der „Teamaufgabe“ unauffällig vom VI ausgetauscht. Dadurch erhielt der MWB 2 eine gegensätzliche „Teamaufgabe“. MWB 1 erhielt die Instruktion zum genauen, aber vor allem schnellen Absolvieren des Pfads. Es handelte sich somit um eine habituell geschwindigkeitsorientierte VP. Wohingegen MWB 2 die Aufgabe erhielt, den Pfad schnell und vor allem genau zu absolvieren. Diese VP war genauigkeitsorientiert ausgeprägt.

### ***Phase 3: Hauptteil***

Nach dieser Anlernphase folgte der Hauptteil der Untersuchung in Abwesenheit des Vls. Bevor der VI den Untersuchungsraum verließ, konnten die MWB Fragen stellen. Zudem wurden die MWB an das Rede- und Gestikenverbot sowie die leistungsabhängige Vergütung erinnert. Als Letztes wurde die Webkamera von dem VI vor den Augen der MWB angeschaltet.

Im Hauptteil der Untersuchung wurden die Solo-Tracking- und kooperativen Tracking-Durchgänge abwechselnd ausgeführt. Nach jedem kooperativen Tracking-Durchgang folgte eine kurze Befragung der MWB (s. Anhang A.2.4). Die Führung der beiden MWB durch die Untersuchung erfolgte mittels der Instruktionen auf dem Computer- respektive Laptopmonitor. Die MWB konnten den Start von jedem Tracking-Durchgang durch Anklicken des Startbuttons am Ende einer Startinstruktion selbst bestimmen.

### ***Phase 4: Abschlussbefragung***

Am Ende der letzten Endinstruktion erfolgte der Hinweis, den VI aufzusuchen. Dieser führte eine mündliche Befragung der MWB durch (s. Anhang A.2.4). Unter anderem wurden das Instruktionsverständnis sowie die Instruktionsbefolgung erfragt. Danach wurden Fragen der MWB zur Untersuchung beantwortet. Die VPn wurden gebeten, in den nächsten Wochen niemandem von dem Inhalt der Untersuchung zu erzählen, so dass alle VPn mit dem gleichen Vorwissen in die Untersuchung gehen. Zuletzt wurde den VPn für deren Teilnahme gedankt und sie erhielten eine Aufwandsentschädigung in Höhe von zehn Euro.

### 3.2.1.2.6 Untersuchungsteilnehmer der Voruntersuchungen

Es folgt die Beschreibung der Akquise und Eigenschaften der VPn

#### 3.2.1.2.6.1 Akquise

Für die Akquise der VPn der Voruntersuchungen wurde der Versuchspersonenserver PESA der Humboldt-Universität zu Berlin verwendet.

Ausgeschlossen wurden Personen, die Links- oder Beidhändigkeit, nicht korrigierte Sehschwäche oder Schwierigkeiten in der Motorik aufwiesen. Ebenso durften die Personen keine Drogen oder Medikamente eingenommen haben, welche einen Einfluss auf die Wahrnehmung oder Psychomotorik haben können. Zudem wurden zur Vermeidung von Alterseffekten in den sensorischen und motorischen Fähigkeiten nur Personen eingeladen, welche älter als 17 und jünger als 60 Jahre alt waren. Zur Vermeidung von sprachlichen Verständnisproblemen wurden nur Personen mit deutscher Muttersprache eingeladen. Des Weiteren durften keine Vorerfahrungen in Bezug auf SAM sowie anderen Tracking-Aufgaben bestehen, um dadurch verursachte unterschiedliche Leistungsniveaus beim Tracking zu vermeiden. Studenten der Psychologie wurden ebenfalls ausgeschlossen, weil es sich bei ihnen i. d. R. nicht um naive VPn handelt. Zudem wurden Personen, welche mindestens einmal als „Nicht-Erscheiner“ bei einer Untersuchung auffielen, nicht eingeladen.

Für die Selektion der VPn hinsichtlich der unspezifischen Strategie wurden alle in PESA in Frage kommenden Personen per E-Mail gebeten, den FESE-R (Nachtwei, 2008b) online auszufüllen. Von diesen wurden dann nur diejenigen zur Teilnahme an den Voruntersuchungen eingeladen, welche bei den beiden Unterskalen „Geschwindigkeits-“ und „Genauigkeitsorientierung“ Extremwerte in Form einer Abweichung von mindestens 10 % vom Mittelwert der Eichstichprobe nach unten oder oben aufweisen (Vorgehen nach Nachtwei, 2008a).

Ursprünglich war geplant, für die drei Voruntersuchungen mindestens je acht MWB-Paare einzusetzen. Dabei war eine Ausbalancierung des Geschlechts in allen drei Voruntersuchungen (vier weibliche und männliche Teams) angedacht. Nachtwei (2008a) hatte in seinen Untersuchungen einen signifikanten Einfluss des Geschlechts auf Zeit und Fehler beim Tracking gefunden. Allerdings zeigte sich die Selektion der VPn nach deren unspezifischen Strategie, erhoben durch den FESE-R (Nachtwei, 2008b), für die drei Voruntersuchungen als nicht versuchsökonomisch. *Personen* mit einer hohen Ausprägung hinsichtlich der *Geschwindigkeitsorientierung* wiesen eine signifikant *höhere Nicht-Erscheinungsrate* auf ( $\chi^2(1) = 9.06$ ,

$p = .003$ ,  $\omega^{13} = .43$ , odds ratio 9.64,  $1 - \beta = .88$ )<sup>14</sup>. Bereits in zwei Untersuchungen konnte Nachtwei (2008a) aufzeigen, dass ein signifikant positiver Zusammenhang zwischen der Genauigkeitsorientierung von VPn und deren Teilnahme-häufigkeit an Experimenten ( $r = .48$ ,  $p_{\text{zweiseitig}} < .01$ ,  $r = .38$ ,  $p_{\text{zweiseitig}} < .05$ ) sowie ein negativer Zusammenhang zwischen der Geschwindigkeitsorientierung von VPn und deren Teilnahme-häufigkeit an Experimenten ( $r = -.44$ ,  $p_{\text{zweiseitig}} < .01$ ) existiert. Ersatzeinladungen von weiteren MWB-Paaren wurden zunächst getätigt. Allerdings war zu diesem Zeitpunkt durch die bisherigen Untersuchungen im ATEO-Projekt der Pool an in Frage kommenden Personen in PESA bereits stark ausgeschöpft (s. Anhang A.2.3). Der normale Zulauf sowie umfangreiche Werbung für PESA durch den Projektpartner und die Autorin erbrachten nur sehr geringfügigen Zuwachs von geeigneten Personen. Statt einer Fortführung der Akquise erschien es sinnvoll, die Notwendigkeit der Selektion von VPn nach deren unspezifischen Strategie für SAM experimentell zu überprüfen (s. Abschnitt 3.2.2), dies vor allem in Hinblick auf die weiteren Untersuchungen sowie der verbliebenden Zeit in der zweiten ATEO-Projektphase. Dadurch konnten je Voruntersuchung nur sechs MWB-Paare gewonnen werden.

### 3.2.1.2.6.2 Beschreibung der Stichproben

Die MWB-Paare waren gleichgeschlechtlich, um Geschlechtereinflüsse auf das Tracking innerhalb eines Untersuchungsdurchgangs zu vermeiden.<sup>15</sup> Dieses Vorgehen hatte sich bereits in der ersten ATEO-Projektphase bewährt. Die Zuordnung der eingeladenen VPn-Teams zu einer der drei Voruntersuchungen erfolgte zufällig, jedoch unter anfänglicher Beachtung der Ausbalancierung des Geschlechts je Voruntersuchung. Bei der Voruntersuchung „Gabelung“ kamen zwei weibliche und vier männliche Teams zum Einsatz. Der Altersdurchschnitt der VPn lag bei 31.17 Jahren ( $SD = 11.19$ ). Für die Voruntersuchung „Hindernis“ wurden je drei Frauen- und Männer-Teams verwendet. In dieser Stichprobe lag das mittlere Alter bei 24.90 Jahren ( $SD = 4.07$ ). Die Wirkung der Manipulation des Joystick-Inputs wurde mittels vier weiblichen und zwei männlichen Teams überprüft. In dieser Voruntersuchung

---

<sup>13</sup> Konventionen für die Effektgröße  $\omega$  nach Cohen (1988, zitiert nach Sedlmeier & Renkewitz, 2008):  $\omega = .1$  ist ein kleiner,  $\omega = .3$  ist ein mittlerer und  $\omega = .5$  ist ein großer Effekt.

<sup>14</sup> Bei dieser Rechnung wurden Personen ausgeschlossen, welche nicht erschienen, jedoch netterweise abgesagt hatten. Insgesamt waren dies fünf Personen, welche allerdings allesamt eine hohe Geschwindigkeitsorientierung aufwiesen. Bei Einschluss dieser fünf Personen ergibt sich ein noch signifikanteres Ergebnis und höhere Effektgröße.

<sup>15</sup> Die zusätzliche Verwendung von gemischten MWB-Paaren zur Überprüfung des tatsächlichen Einflusses dieser Geschlechtskonstellation auf das Tracking stand nicht im Fokus der Forschung. Zudem hätte eine zusätzliche Verwendung von gemischten MWB-Paaren hinsichtlich des Geschlechts zu einer bis zwei weiteren Stufen (Je nachdem, ob für das Geschlecht die Zuordnung zur Rolle des MWB 1 und 2 berücksichtigt wird.) der kontrollierten Variable „Geschlecht“ geführt. Der Verzicht auf gemischte MWB-Teams aus diesen Gründen gilt für alle weiteren Untersuchungen im Rahmen des ATEO-Projekts.

waren die VPn im Durchschnitt 27.75 Jahre alt ( $SD = 8.37$ ). Die drei Stichproben bestanden hauptsächlich aus Studenten ( $H = 20$ ) und Berufstätigen ( $H = 11$ ). In der Voruntersuchung „Hindernis“ war der größte Anteil an Studenten ( $H = 9$ ) vertreten. Die Berufs- und Studienrichtungen weisen eine hohe Spannweite auf. Bis auf das Studium der Linguistik und Biologie (jeweils  $H = 2$ ) kam keine Richtung mehrfach vor. Details zu den Stichproben sind dem Anhang A.3.1 zu entnehmen.

### **3.2.1.3 Ergebnisse**

In diesem Kapitel wird die Aufbereitung der erhobenen objektiven Daten während des Trackings der MWB und die sich daran anschließende Analyse beschrieben. Die Ergebnisse werden in der Reihenfolge der drei Voruntersuchungen vorgestellt.

#### **3.2.1.3.1 Aufbereitung der objektiven Tracking-Daten**

Pro Tracking-Durchgang wurden die relevanten Untersuchungsdaten ca. alle 39 ms gemessen und als eine Datenzeile in ein XML-Logfile geschrieben. Die sich so ergebende Datenmenge wurde mit einem eigens von einem Projektmitarbeiter für die Voruntersuchung angefertigten Datenaufbereitungs-Tool für die relevanten Pfadabschnitte aufbereitet und in XLS-Dateien konvertiert (s. Abschnitt 3.2.1.3.2). Diese Dateien wurden von der Autorin manuell in eine XLS-Datei zusammengefügt. Das Tool ermittelt u. a. die Summe der vertikalen und horizontalen Joystick-Auslenkungen pro MWB sowie des NrSensors, die benötigte Tracking-Zeit und die Häufigkeit der Wahl der beiden Pfadoptionen je Gabelung.

#### **3.2.1.3.2 Analyse der objektiven Tracking-Daten**

Mittels der drei Voruntersuchungen sollte die Forschungsfragestellung 1 aufgeklärt werden, welche zwei der vier Varianten einer jeden uV am geeignetsten für den Einsatz in SAM 2.0 sind. Die Varianten sollten der Erhöhung der spezifisch auf Konflikte der MWB bezogenen wahrnehmbaren Komplexität von SAM dienen. Aufgrund der geringen Menge an gewonnenen Teams wurden nur deskriptive Analysen zur Aufklärung dieser Forschungsfragestellung durchgeführt.

Für die Entscheidung, welche zwei Gabelungen am geeignetsten zur Komplexitätserhöhung von SAM sind, wurde das Verhalten der MWB bei der Objektregelung unmittelbar vor und im Bereich der Gabelungen betrachtet. Der Bereich vor der Gabelung war von Interesse, weil in diesem die Entscheidung durch die MWB für eine der beiden Pfadoptionen gefällt werden muss. Bei der Wahl der Pfadoptionen sollte es aufgrund der unterschiedlichen Strategien zu einem Entscheidungskonflikt zwischen den MWBn kommen, welcher sich in

unterschiedlichen Joystick-Auslenkungen der MWB und einer dadurch verursachten höheren Fehlerrate widerspiegeln sollte. Deshalb wurden für den Bereich vor der Gabelung die Differenzen der Joystick-Auslenkungen der MWB sowie die Fehleranzahl betrachtet. Je höher die Werte diesbzgl. waren, desto geeigneter erschien eine Gabelung zur Erhöhung der durch Operateure und Entwickler wahrnehmbaren Komplexität von SAM. Der Eintritt in eine Gabelung, deren Kurvenbereich respektive schmalerer Pfad sollten zu mehr Fehlern führen. Es wurden jedoch keine extremen Unterschiede zwischen den Joystick-Auslenkungen der beiden MWB im Gabelungsbereich erwartet, da keine Entscheidungssituation vorkommt. Am geeignetsten für die weitere Verwendung in SAM sind die Gabelungsvarianten, welche die beiden höchsten Fehlerwerte im Bereich der Gabelungen aufweisen. Da die Wahl einer Pfadoption durch die Tracking-Strategien beeinflusst sein sollte, wird für den Bereich der Gabelung zudem die Häufigkeit der Objektregelung auf der einen oder der anderen Pfadoption betrachtet. Dabei sollte ein ausgewogenes Verhältnis zwischen der Wahl der Pfadoptionen bestehen. Für die Operateure oder Entwickler entspricht dies einem zufälligen Anteil des Prozessverlaufs von SAM.

Die Bereiche vor, während und nach dem Hindernis-Slalom wurden analysiert, um die zwei am besten geeigneten statischen Hindernisse zur Erhöhung der Komplexität von SAM auswählen zu können. In den Bereichen vor und während der Hindernisse sollten die gegensätzlichen Tracking-Strategien zu einer unterschiedlichen Objektregelung führen, welche noch in den Bereich nach den Hindernissen hineinreicht. Somit sollten die beiden geeignetsten Hindernisse in diesen drei Bereichen bzgl. der Differenzen der Joystick-Auslenkungen der MWB sowie der Fehlerrate die höchsten Werte aufweisen.

Für die Selektion der zwei geeignetsten Manipulationen des Joystick-Inputs wurde der Bereich der Manipulation selbst sowie die sich direkt danach anschließende Phase betrachtet. Die Manipulation des Joystick-Inputs sollte zu der subjektiven Wahrnehmung eines Leistungsunterschieds zwischen den MWBn führen, welche nach dem 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) einen Konflikt zwischen diesen verursacht. Dieser sollte sich in gegensätzlichen Joystick-Bewegungen sowie einer erhöhten Fehlerrate zeigen. Es wird vermutet, dass sich der Einfluss der Manipulation auf das Verhalten der MWB in den Bereich nach der Manipulation überträgt. Die geeignetsten Manipulationen des Joystick-Inputs sollten die höchsten Differenzen der Joystick-Auslenkungen der MWB sowie die höchste Fehlerrate aufweisen.

Alle Bereiche sollten zudem schnell absolviert werden, um von einem Operateur oder den Automaten schnelle Reaktionen abzuverlangen. Des Weiteren sollten die Werte kleine Streuungen aufweisen, wodurch von einer höheren Generalität dieser ausgegangen werden kann.

Die Bereiche entsprechen 800 Pixeln auf der Vertikalen. Einzig der Bereich während eines Hindernisses umfasst 300 Pixel.

### 3.2.1.3.3 Ergebnisse der Analysen zur Forschungsfragestellung 1

Nachfolgend werden die Ergebnisse der deskriptiven Analysen zur Aufklärung der Forschungsfragestellung 1 vorgestellt. Dabei werden die Ergebnisse der vier Bedingungen einer uV je Bereich in der Reihenfolge des erreichten Rangplatzes (s. auch jeweilige Tabelle) hinsichtlich der Erfüllung der gewünschten Ausprägungen der aVn beschrieben.

#### 3.2.1.3.3.1 Gabelungen

Tabelle 6 stellt die erhobenen Daten der aVn für den Bereich vor den Gabelungen dar.

*Tabelle 6: Mittelwerte und Standardabweichungen der Differenzen bei der horizontalen und vertikalen Joystick-Auslenkung der MWB, der benötigten Zeit und der Fehler beim Objekt-Tracking im Bereich vor der Gabelung*

Rangplatz	Bedingung	Differenz horizontale Joystick-Auslenkung MWB 1 und 2		Differenz vertikale Joystick-Auslenkung MWB 1 und 2		benötigte Zeit in ms		Fehler	
		M	SD	M	SD	M	SD	M	SD
4	<i>ell</i>	<b>8306</b>	17816	6446	31173	1698	725	12	<b>17</b>
1	<i>elr</i>	<b>8710</b>	<b>11687</b>	11714	<b>20811</b>	<b>1619</b>	<b>590</b>	<b>16</b>	28
2	<i>rll</i>	5210	13912	<b>25728</b>	33847	<b>1630</b>	605	<b>26</b>	26
3	<i>rlr</i>	6877	<b>11794</b>	<b>27048</b>	<b>23850</b>	1651	<b>599</b>	11	<b>16</b>

*Anmerkungen.* M = Mittelwert; SD = Standardabweichung; Bedingung Gabelung: rll = runde, lange Pfadoption nach links ausgerichtet; rlr = runde, lange Pfadoption nach rechts ausgerichtet; ell = eckige, lange Pfadoption nach links ausgerichtet; elr = eckige, lange Pfadoption nach rechts ausgerichtet; die beiden geeignetsten Ausprägungen einer aV sind fett markiert

Die Gabelung „*elr*“ weist die höchsten durchschnittlichen Werte bei der Differenz der horizontalen Joystick-Auslenkung zwischen den MWBn und der Geschwindigkeit mit jeweils der kleinsten Streuung vor. Des Weiteren wurde vor dieser Gabelung im Mittel die zweithöchste Fehleranzahl produziert. Die Gabelung „*rll*“ besitzt die zweithöchsten Werte bzgl. der durchschnittlichen Differenz zwischen den MWBn bei der vertikalen Objektregelung und der Geschwindigkeit. Bei dieser Gabelung zeigte sich der höchste mittlere Fehlerwert. Bei der Gabelung „*rlr*“ bewegten MWB 1 und 2 im Durchschnitt den Joystick auf der Vertikalen am unterschiedlichsten. Die Gabelung „*ell*“ verursachte im Mittel den zweithöchsten Differenzwert für die horizontale Joystick-Auslenkung der beiden MWB.

Im Bereich der Gabelungen liegen bei „*rll*“ die zweithöchsten Werte hinsichtlich der benötigten Zeit und Fehleranzahl vor (s. Tabelle 7). Bei der Gabelung „*elr*“ wurden im Durch-

schnitt die meisten Fehler produziert. Zudem weisen diese beiden Gabelungen die ausgewogensten Verhältnisse bei der Wahl zwischen den Pfadoptionen auf. Die schnellste Tracking-Zeit konnte bei der Gabelung „**rlr**“ erreicht werden.

Tabelle 7: Häufigkeiten und Prozente der Gabelungspfadwahl sowie Mittelwerte und Standardabweichungen der benötigten Zeit und der Fehler beim Objekt-Tracking im Bereich innerhalb der Gabelung

Rangplatz	Bedingung	Wahl der Pfadoption				benötigte Zeit in ms		Fehler	
		Pfad gerade, dünn (schnell)		Pfad eckig/rund, breit (genau)		M	SD	M	SD
		Häufigkeit	Prozent	Häufigkeit	Prozent				
4	<i>ell</i>	4	22	14	78	3499	761	67	<b>43</b>
2	<i>elr</i>	8	<b>44</b>	10	<b>56</b>	3590	787	<b>81</b>	61
1	<i>rl</i>	7	<b>39</b>	11	<b>61</b>	<b>3201</b>	<b>484</b>	<b>73</b>	<b>51</b>
3	<i>rlr</i>	5	28	13	72	<b>3086</b>	<b>567</b>	72	65

Anmerkungen. M = Mittelwert; SD = Standardabweichung; Bedingung Gabelung: rll = runde, lange Pfadoption nach links ausgerichtet; rlr = runde, lange Pfadoption nach rechts ausgerichtet; ell = eckige, lange Pfadoption nach links ausgerichtet; elr = eckige, lange Pfadoption nach rechts ausgerichtet; die beiden geeignetsten Ausprägungen einer aV sind fett markiert

Aufgrund der Werte der aVn wurden für die Gabelungen Rangplätze hinsichtlich der Eignung zur Erhöhung der spezifisch auf Konflikte der MWB bezogenen wahrnehmbaren Komplexität von SAM bestimmt (s. Tabelle 6 und Tabelle 7). Aus den Rangplätzen ist ersichtlich, dass in Summe die Werte der aVn bei den Gabelungen „rll“ und „elr“ am ehesten die präferierten Ausprägungen aufweisen.

### 3.2.1.3.3.2 Statische Hindernisse

Tabelle 8 präsentiert die erhobenen Daten der aVn für die Bereiche vor, innerhalb und nach den Hindernissen. Im Bereich vor den Hindernissen liegen für das Hindernis „**25lr**“ bzgl. der durchschnittlichen Differenz der Joystick-Auslenkung zwischen den beiden MWBn die zweithöchsten Werte auf der Horizontalen und die höchsten Werte auf der Vertikalen vor. Wobei nur letzterer Wert die kleinste Streuung besitzt. Zudem produzierten die MWB vor diesem Hindernis im Mittel die kürzeste Tracking-Zeit mit der kleinsten Streuung sowie der zweitgrößten Fehleranzahl. Das Hindernis „**50lr**“ weist in diesem Bereich den durchschnittlich höchsten horizontalen und den zweithöchsten vertikalen Differenzwert der Joystick-Auslenkung der beiden MWB auf. Der Bereich vor dem Hindernis „**25rl**“ kann im Durchschnitt am zweitschnellsten absolviert werden.

Im Bereich der Hindernisse besitzt erneut „**25lr**“ bei den aVn, bis auf die Variable „Fehler“, durchschnittlich die zweit- oder höchsten Werte. Bei dem Hindernis „**50lr**“ zeigte

sich wieder im Mittel die höchste Differenz der horizontalen Joystick-Auslenkung der MWB. Zudem verursachten die MWB beim Passieren dieses Hindernisses abermals durchschnittlich die meisten Fehler, diesmal mit der zweitkleinsten Streuung. Jedoch weist nun das Hindernis „**50rl**“ hinsichtlich der vertikalen Joystick-Auslenkung der MWB die zweithöchste mittlere Differenz auf. Ferner wurden im Durchschnitt bei diesem Hindernis am zweithäufigsten Fehler produziert. Am Hindernis „**25rl**“ konnte das Objekt im Mittel am schnellsten vorbei geregelt werden, jedoch ergab sich dabei die kleinste Fehleranzahl. Bei beiden Werten liegt die kleinste Streuung vor.

Tabelle 8: Mittelwerte und Standardabweichungen der Differenzen bei der horizontalen und vertikalen Joystick-Auslenkung der MWB, der benötigten Zeit und der Fehler beim Objekt-Tracking beim Einsatz von Hindernissen

		Differenz horizontale Joystick-Auslenkung MWB 1 und 2		Differenz vertikale Joystick-Auslenkung MWB 1 und 2		benötigte Zeit in ms		Fehler	
Rangplatz	Bedingung	M	SD	M	SD	M	SD	M	SD
im Bereich vor dem Hindernis									
1	25lr	17573	21634	13244	20697	1974	237	45	27.20
3	25rl	12572	15908	504	25695	2035	245	28	19.90
2	50lr	20915	20807	6391	35439	2126	408	48	33.90
4	50rl	14661	16526	183	32541	2096	414	35	26.60
im Bereich innerhalb des Hindernisses									
1	25lr	49430	49846	57623	45775	4873	677	131	57.38
4	25rl	29340	42741	32754	43396	4621	497	80	40.68
2	50lr	58533	45694	35252	95243	5142	1011	188	42.86
3	50rl	48045	35370	54609	76388	4942	943	152	63.25
im Bereich nach dem Hindernis									
1	25lr	13673	15910	18126	21359	1543	649	36	42.50
4	25rl	4982	15996	4622	20000	1523	728	25	24.10
3	50lr	14603	25862	4015	33453	1579	778	32	24.50
2	50rl	12671	12071	11318	17029	1441	616	30	23.30

Anmerkungen. M = Mittelwert; SD = Standardabweichung; Bedingung statisches Hindernis: 25lr = erst links, dann rechts 25 % Pfadabdeckung; 25rl = erst rechts, dann links 25 % Pfadabdeckung; 50lr = erst links, dann rechts 50 % Pfadabdeckung; 50rl = erst rechts, dann links 50 % Pfadabdeckung; die beiden geeignetsten Ausprägungen einer aV sind fett markiert

Im Bereich nach den Hindernissen weist „**25lr**“ weiterhin bei den aVn, bis auf die „benötigte Zeit“, durchschnittlich die zweit- oder höchsten Werte auf. Bei dem Hindernis „**50rl**“ liegt im Durchschnitt wieder der zweithöchste Differenzwert bzgl. der vertikalen Joystick-Auslenkung zwischen den MWBn und erstmalig die höchste Geschwindigkeit vor. Bei



dem Hindernis „**50lr**“ zeigten die MWB erneut im Mittel bei der Bewegung des Joysticks auf der Horizontalen die größten Unterschiede. Zudem weist dieses Hindernis diesmal nur den zweithöchsten durchschnittlichen Fehlerwert auf. Das Hindernis „**25rl**“ konnte lediglich bei der benötigten mittleren Zeit den zweitkleinsten Wert erreichen.

Zusammenfassend lässt sich feststellen, dass über die drei betrachteten Bereiche hinweg bei den beiden Hindernissen „25lr“ und „50lr“ am ehesten das erwünschte Wertemuster der aVn vorliegt (s. Rangplätze Tabelle 8).

### 3.2.1.3.3 Manipulation des Joystick-Inputs

Tabelle 9 zeigt die erhobenen Daten der aVn für die Bereiche während und nach der Joystick-Manipulation. Während der Reduzierung des prozentualen Anteils des Joystick-Inputs an der Objektregelung von einem der beiden MWB um **8 %** ergaben sich im Durchschnitt die höchsten Werte hinsichtlich des Unterschieds der vertikalen Joystick-Auslenkung zwischen den MWBn und der Fehleranzahl. Ebenso zeigte sich die zweitkürzeste Tracking-Zeit. Bei der Bedingung „**4 %**“ liegen im Mittel die höchsten Werte bei der Differenz der horizontalen Joystick-Auslenkung zwischen den MWBn sowie der Geschwindigkeit vor. Für die Variante „**16 %**“ konnten die zweithöchsten Werte bzgl. der Differenz der vertikalen Joystick-Auslenkung zwischen den MWBn und der Fehleranzahl gefunden werden. Unter der Bedingung „**12 %**“ produzierten die beiden MWB bei der horizontalen Joystick-Auslenkung den zweithöchsten Wert.

Für den Bereich nach der Manipulation des Joystick-Inputs zeigten sich im Durchschnitt unter der Bedingung „**8 %**“ erneut die höchsten Werte bei der Differenz der vertikalen Joystick-Auslenkung zwischen den MWBn und der Fehleranzahl. Zudem weist diese Bedingung diesmal im Mittel den zweithöchsten Wert bzgl. der Differenz der horizontalen Joystick-Auslenkung zwischen den MWBn auf. Die Variante „**12 %**“ erreichte bei der Differenz der horizontalen Joystick-Auslenkung zwischen den MWBn den höchsten Mittelwert mit der kleinsten Streuung. Zudem liegt bei dieser im Mittel die geringste Zeit vor. Bei der Bedingung „**4 %**“ finden sich durchschnittlich die zweithöchsten Werte bei der Differenz der vertikalen Joystick-Auslenkung zwischen den MWBn und der Geschwindigkeit. Durch die Reduzierung des prozentualen Anteils des Joystick-Inputs von einem der beiden MWB an der Objektregelung um **16 %** ergab sich im Durchschnitt der zweithöchste Fehlerwert.

Die Werte der aVn über die beiden Bereiche entsprechen bei der Manipulation des Joystick-Inputs um **8 %** am ehesten dem präferierten Wertemuster (s. Rangplatz Tabelle 9). Bei den anderen drei Bedingungen ist augenscheinlich kein Unterschied hinsichtlich der Eignung

zur Erhöhung der spezifisch auf Konflikte der MWB bezogenen wahrnehmbaren Komplexität von SAM festzustellen.

Tabelle 9: Mittelwerte und Standardabweichungen der Differenzen bei der horizontalen und vertikalen Joystick-Bewegung der MWB, der benötigten Zeit und Fehler beim Objekt-Tracking beim Einsatz der Joystick-Manipulation

		Differenz horizontale Joystick-Auslenkung MWB 1 und 2		Differenz vertikale Joystick-Auslenkung MWB 1 und 2		benötigte Zeit in ms		Fehler	
Rang-platz	Bedingung	M	SD	M	SD	M	SD	M	SD
im Bereich der Joystick-Manipulation									
2	4 %	62744	74603	2047	123442	6582	1443	150.75	78.10
1	8 %	47832	54164	30021	163463	6809	1561	159.08	105.60
4	12 %	49406	64389	6584	140313	6933	1694	120.50	72.30
3	16 %	28214	46496	14493	102425	7193	1365	151.00	106.10
im Bereich nach der Joystick-Manipulation									
3	4 %	7949	25190	3787	30056	1702	359	40	32
1	8 %	10920	19101	12084	43437	1750	340	68	53
2	12 %	11796	16680	2051	32086	1699	342	37	32
4	16 %	3510	16869	381	26737	1703	231	57	37

Anmerkungen. M = Mittelwert; SD = Standardabweichung; Bedingung Input-Manipulation in %: 4, 8, 12, 16; die beiden geeignetsten Ausprägungen einer aV sind fett markiert

### 3.2.1.4 Diskussion und Interpretation der Ergebnisse zur Forschungsfragestellung 1

Die Werte der aVn bei den **Gabelungen** „*rl*“ und „*el*“ entsprechen am ehesten den gewünschten Ausprägungen (höchste Differenz bei der horizontalen und vertikalen Joystick-Auslenkung zwischen den MWBn, höchste Fehleranzahl, kürzeste Tracking-Zeit sowie ausgewogenste Pfadauswahl). Daher wurden diese beiden Gabelungen als am geeignetsten bewertet, für den Einsatz in der nächsten SAM-Version 2.0. Die Gabelungen weisen das Potential auf, die durch einen Operateur oder Entwickler wahrnehmbare Komplexität von SAM zu erhöhen. Dies zeigt sich zum einen in der erhöhten Fehlerrate bei Eintritt in die Gabelung sowie der Wahl beider Pfadoptioenen einer Gabelung durch die MWB.

Bei den **Hindernissen** liegt „*25lr*“ in allen Bereichen auf Rangplatz eins, da deren Werte bei den aVn am ehesten das präferierte Wertemuster (höchste Differenz bei der horizontalen und vertikalen Joystick-Auslenkung zwischen den MWBn, höchste Fehleranzahl sowie kürzeste Tracking-Zeit) aufweisen. Daneben besteht beim Hindernis „*50lr*“ über die drei Bereiche hinweg das zweitbeste Wertemuster der aVn. Daher wurden diese beiden Hindernisse in die nächste SAM-Version 2.0 integriert. Die Hindernisse besitzen das Potential, die durch einen Operateur oder Entwickler wahrnehmbare Komplexität von SAM zu erhöhen. Dies zeigt

sich in den höheren Werten der aVn beim Passieren der vier Varianten im Vergleich zu den Bereichen vor und nach den Hindernissen.

Bei der *Manipulation des Joystick-Inputs* fiel die Entscheidung, aufgrund der gefundenen Datenlage, schwerer. Am ehesten weist die Bedingung „8 %“ das gewünschte Ausprägungsmuster (höchste Differenz bei der horizontalen und vertikalen Joystick-Auslenkung zwischen den MWBn, höchste Fehleranzahl sowie kürzeste Tracking-Zeit) bei den aVn auf. Allerdings besitzen diese Werte die zweit- und höchsten Streuungen. Von den anderen drei Bedingungen sticht keine hinsichtlich der gewünschten Ausprägungen der aVn hervor. Im Zuge der, von Projektmitgliedern und der Autorin parallel zu diesen Voruntersuchungen durchgeführten, Hierarchischen Aufgabenanalysen (HAA) zeigte sich, dass die Reduzierung des prozentualen Anteils des Joystick-Inputs von einem oder gar beiden MWBn an der Objektregelung eine Funktionalität des Operatorsarbeitsplatzes sowie der Automaten sein kann (s. Abschnitt 3.3.1.1). Aufgrund der vorgefundenen, augenscheinlich nicht eindeutigen Datenlage und den Ergebnissen der HAA wurde beschlossen, die Manipulation des Joystick-Inputs nicht weiter zur Steigerung der Komplexität von SAM zu verfolgen. Dadurch sollte den Operateuren und den Automaten die Möglichkeit gegeben werden, die Führung des SAM-Prozesses durch Veränderung des Input-Verhältnisses zwischen den MWBn auszuüben.

Die Analyse der Joystick-Auslenkungen der beiden MWB zeigte auf, dass sie eine unterschiedliche Objektregelung ausgeführt haben. Im Bereich der Gabelungen und Hindernisse weisen die MWB zumeist gegensätzliche Auslenkungen auf. Wobei der genauigkeitsorientierte MWB immer eine stärkere Verringerung der Objektgeschwindigkeit anstrebt. Bei den horizontalen Joystick-Auslenkungen finden sich in den meisten Bereichen gegensätzliche oder zumindest unterschiedlich intensive Auslenkungen zwischen den MWBn, wobei der Geschwindigkeitsorientierte die stärkere Auslenkung aufweist.<sup>16</sup> Details s. Anhang A.3.2. Diese unterschiedlichen Auslenkungen sollten zur Wahrnehmung eines Leistungsunterschieds zwischen den MWB führen. Dies sollte wiederum nach dem 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) eine Erhöhung des Konflikts zwischen ihnen verursachen.

### 3.2.1.5 Fazit und Kritik

Zusammenfassend erscheinen, aufgrund der gewonnenen Daten bzgl. der aVn in den drei Voruntersuchungen, die Gabelungen „rll“ und „elr“ sowie die Hindernisse „25lr“ und

---

<sup>16</sup> Eine genaue, zwischen dem Verhalten der MWB vergleichende, Analyse der horizontalen Joystickdaten fällt ohne Bezug zum Pfad schwer. Die vorhandenen Auslenkungsdaten der Joysticks aus den Voruntersuchungen können nur über einen aufwändigen, teilweise manuellen Analyseweg in Relation zum Pfad gesetzt werden. Deshalb wurde im Rahmen dieser Auswertung darauf verzichtet.

„50lr“ am geeignetsten zu sein, für die weitere Verwendung in SAM 2.0 zur Erhöhung der durch einen Operateur oder Entwickler wahrnehmbaren Komplexität. Alle anderen Varianten inklusive der vier Manipulationen der Joystick-Inputs der MWB wurden nicht in die nachfolgende Version SAM 2.0 implementiert. Es muss kritisch angemerkt werden, dass diese Entscheidungen auf sehr geringen Datenmengen (je uV sechs MWB-Teams) beruhen. Dadurch können Ausreißer die Ergebnisse stärker als in großen Stichproben beeinflusst haben. Zudem ist die Selektion der geeigneten Varianten der uVn aufgrund der rein deskriptiven Analyse interpretativ. Die Entscheidungen wurden deshalb von der Autorin im ATEO-Team diskutiert und von diesem unterstützt.

### **3.2.2 Voruntersuchung: Vorselektion der MWB**

Aufgrund des stark ausgeschöpften Pools an in Frage kommenden Personen in PESA für die Rolle von einem der MWB, wurde in einer vierten Voruntersuchung der Effekt der Vorselektion der VPn hinsichtlich deren unspezifischen Strategie auf deren Objektregelungsverhalten ermittelt. Dafür wurde erneut die Voruntersuchung „Gabelung“ durchgeführt. Diesmal werden die MWB jedoch nicht nach deren unspezifischen Strategie vorselektiert. Die Tracking-Leistung der MWB dieser Voruntersuchung wurde mit der der Voruntersuchung „Gabelung mit Vorselektion der der MWB nach deren unspezifischen Strategie“ (s. Abschnitt 3.2.1) verglichen.

Der Untersuchungsort, die Erhebungsinstrumente, das Untersuchungsmaterial, der Aufbau und Ablauf dieser vierten Voruntersuchung waren identisch mit denen der drei ersten Voruntersuchungen (s. Abschnitt 3.2.1.2). Die Operationalisierung der aVn (Differenz bei der horizontalen und vertikalen Joystick-Auslenkung zwischen den MWBn, Fehler sowie benötigte Zeit beim Tracking) entsprachen denen der anderen drei Voruntersuchungen. Jedoch lag diesmal der untersuchungsrelevante Bereich nicht nur rund um die neuen SAM-Elemente, sondern erstreckte sich vom Start bis zum Ziel eines Tracking-Pfads. Ebenso wurden die gleichen Kontrollvariablen erhoben und gemeinsam mit den Daten aus den anderen drei Voruntersuchungen analysiert. Die Details (Erhebungszeitpunkt, Analyse, Ergebnisse) dazu sind, aus Platzgründen, für den interessierten Leser im Anhang B.3.1 beschrieben. Um Redundanzen zu vermeiden, werden im Folgenden nur Sachverhalte erwähnt, welche sich von der Voruntersuchung „Gabelung mit Vorselektion der MWB nach deren unspezifischen Strategie“ unterscheiden.

#### **3.2.2.1 Fragestellung und Hypothesen**

Dieser Abschnitt widmet sich der Herleitung der Forschungsfragestellung 2, die dem Vergleich der zwei Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach

deren unspezifischen Strategie“ zugrunde liegt. Hierbei handelt es sich um eine methodische Fragestellung hinsichtlich der Optimierung der Versuchsumgebung zu einem versuchsökonomischeren Instrument. Diese stellt keine inhaltliche Forschungsfragestellung dieser Dissertation oder des ATEO-Projekts dar. Ferner werden die daraus abgeleiteten Hypothesen vorgestellt. Bei diesen handelt es sich um Gleichheitshypothesen.

### **3.2.2.1.1 Herleitung der zweiten Forschungsfragestellung**

Ausgehend von der Gültigkeit des 3K-Modells (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) wurden VPn zur Bildung eines MWB-Paares für SAM bisher hinsichtlich deren unspezifischen Strategie so ausgewählt, dass diese gegensätzliche Ausprägungen (schnell vs. genau zu handeln) aufweisen. Die Ausprägung der jeweiligen unspezifischen Strategie wird in SAM durch eine dazu passende spezifische Strategie in Form einer Instruktion für die Tracking-Aufgabe (schnell vs. genau zu agieren) unterstützt. Durch die konträren Ausprägungen bei der unspezifischen als auch spezifischen Strategie der MWB sollten diese unterschiedliche Tracking-Ziele aufweisen, welche sich in unterschiedlichen Objektregelungsverhalten widerspiegeln. Die Wahrnehmung dieser unterschiedlichen Verhaltensweisen sollte wiederum einen Konflikt zwischen den MWBn verursachen. Der Konflikt seinerseits sollte das unterschiedliche Objektregelungsverhalten der MWB verstärken. Die unspezifische Strategie sollte sich indirekt und die spezifische Strategie direkt auf die kooperative Tracking-Leistung auswirken. (Nachtwei, 2008a)

Empirische Belege für die Beeinflussung der kooperativen Tracking-Leistung durch die unspezifische als auch spezifische Strategie der MWB existierten bis dato nicht. Da jedoch eine ausreichende Anzahl an VPn mit den notwendigen Extremausprägungen bei der unspezifischen Strategie für die noch anstehenden Untersuchungen im Rahmen des ATEO-Projekts nur mit sehr hohem Zeit- und Kostenaufwand akquiriert hätten werden können, entstand folgende Fragestellung:

#### **Forschungsfragestellung 2:**

Stellt der indirekte Einfluss der unspezifischen Strategie auf das Tracking-Verhalten der beiden MWB einen Mehrwert zu dem direkten Einfluss der spezifischen Strategie dar?

Nachfolgend werden die aus der Forschungsfragestellung abgeleiteten Hypothesen vorgestellt:

## **2.1 Forschungshypothesen bezüglich $\alpha V$ „Konflikt zwischen MWBn beim Objekt-Tracking“**

Um bei den zukünftigen Untersuchungen mit SAM auf die Vorselektion von VPn nach deren unspezifischen Strategie verzichten zu können, darf diese keinen über den der spezifischen Strategie hinausgehenden Einfluss auf die Konflikte der MWB haben. Diese Konflikte spiegeln sich in einem unterschiedlichen Objektregelungsverhalten der MWB wider. Somit sollte kein Unterschied zwischen den beiden Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“ bzgl. der Differenz der horizontalen und vertikalen Joystick-Auslenkung der beiden MWB beim kooperativen Tracking bestehen. Daher wurden zur Aufklärung der Forschungsfragestellung 2 die folgenden zwei Hypothesen aufgestellt:

*F.2.1.1: Es existiert kein Unterschied bei der Differenz der horizontalen Joystick-Auslenkung der beiden MWB beim Tracking zwischen den beiden Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“.*

*F.2.1.2: Es existiert kein Unterschied bei der Differenz der vertikalen Joystick-Auslenkung der beiden MWB beim Tracking zwischen den beiden Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“.*

## **2.2 Forschungshypothese bezüglich $\alpha V$ „Genauigkeit beim Objekt-Tracking“**

Wenn sich das Objektregelungsverhalten der MWB zwischen den beiden Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“ nicht differenziert, sollten auch keine Unterschiede hinsichtlich der Genauigkeit beim kooperativen Tracking der MWB vorliegen. Daraus ergab sich zudem die nachfolgende Hypothese:

*F.2.2: Es existiert kein Unterschied bei der Anzahl der Fehler beim Tracking zwischen den beiden Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“.*

## **2.3 Forschungshypothese bezüglich $\alpha V$ „Geschwindigkeit beim Objekt-Tracking“**

Ebenso sollte kein Unterschied hinsichtlich der Geschwindigkeit beim kooperativen Tracking der MWB existieren. Daher wurde die folgende Hypothese aufgestellt:

*F.2.3: Es existiert kein Unterschied bei der benötigten Zeit zum Tracking zwischen den beiden Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“.*

### 3.2.2.2 Methoden

#### 3.2.2.2.1 Operationalisierung der unabhängigen Variable

##### *Voruntersuchung 4: uV „Vorselektion“*

Diese uV bestand aus zwei Stufen, den beiden Stichproben aus den zwei Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“:

*Vorselektion 1: Gabelung, MWB vorselektiert nach unspezifischer Strategie (Gs)*

*Vorselektion 2: Gabelung, MWB nicht vorselektiert nach unspezifischer Strategie (Gu)*

Es handelte sich somit um einen Between-Faktor bzw. Gruppierungsfaktor.

#### 3.2.2.2.2 Experimentelles Design

Die Anzahl und Art der Tracking-Durchgänge der Voruntersuchung „Gabelung ohne Vorselektion der MWB“ entsprachen denen der Voruntersuchung zur „Gabelung mit Vorselektion der MWB nach deren unspezifischen Strategie“ (s. Abschnitt 3.2.1). Das experimentelle Design bestand aus einem einfaktoriellen Versuchsplan. Der Gruppierungsfaktor „Vorselektion“ war zweifach gestuft (vorselektiert vs. nicht-vorselektierte MWB). Somit handelte es sich um ein Between-Groups-Design.

#### 3.2.2.2.3 Durchführung der Voruntersuchung

Der Untersuchungsort, -aufbau, -materialien sowie die zeitliche Planung entsprachen den ersten drei Voruntersuchungen zu SAM 1.5 (s. Abschnitt 3.2.1.2.5). Diese hatten sich dort bereits bewährt, so dass kein Probedurchlauf notwendig war.

##### 3.2.2.2.3.1 Ablauf der Voruntersuchung

Der Ablauf der Voruntersuchung „Gabelung ohne Vorselektion der MWB“ unterschied sich nur in einem Sachverhalt von dem der anderen drei Voruntersuchungen (s. Abschnitt 3.2.1.2.5.2). Die Zuordnung der Rollen MWB 1 und 2 erfolgte diesmal zufällig. Die VP, welche links vor dem Monitor saß, schlüpfte in die Rolle von MWB 1 und die VP, welche sich rechts davon platzierte, übernahm die Rolle von MWB 2.

#### 3.2.2.2.4 Untersuchungsteilnehmer

Es folgt die Beschreibung der Akquise und der Eigenschaften der VPn der vierten Voruntersuchung „Gabelung ohne Vorselektion der MWB“.

#### **3.2.2.2.4.1 Akquise**

Das Vorgehen bei der Akquise von VPn für die Voruntersuchung „Gabelung ohne Vorselektion der MWB“ stimmte ebenfalls mit dem der anderen drei Voruntersuchungen zu SAM 1.5 überein (s. Abschnitt 3.2.1.2.6.1). Einzig die Vorselektion der VPn nach deren unspezifischen Strategie mit Hilfe des FESE-R's (Nachtwei, 2008b) wurde nicht vorgenommen.

#### **3.2.2.2.4.2 Beschreibung der Stichprobe**

Die MWB-Paare waren gleichgeschlechtlich. Für die Stichprobe mit den nicht nach unspezifischer Strategie vorselektierten VPn kamen ein weibliches und fünf männliche Teams zum Einsatz. In dieser Stichprobe lag das mittlere Alter bei 23.90 Jahren ( $SD = 2.07$ ). Bis auf zwei Studenten der Ingenieurwissenschaft bestand die Stichprobe aus Studenten der unterschiedlichsten Fachrichtungen (s. Anhang A.3.1). Die Zusammensetzung der Stichprobe mit den nach unspezifischer Strategie vorselektierten VPn kann im Abschnitt 3.2.1.2.6.2 nach gelesen werden.

#### **3.2.2.3 Ergebnisse**

In diesem Abschnitt werden die Aufbereitungen und die sich anschließenden Analysen der erhobenen Daten während des Trackings durch die MWB vorgestellt. Die Ergebnisse werden in der Reihenfolge der Hypothesen präsentiert.

##### **3.2.2.3.1 Aufbereitung der objektiven Tracking-Daten**

Die Aufbereitung der objektiven Tracking-Daten entsprach dem unter Abschnitt 3.2.1.3.1 beschriebenen Vorgehen. Jedoch sind die erhobenen Rohdaten nicht für die einzelnen Pfadbereiche, sondern für die vier Tracking-Pfade mit Gabelungen vom Start bis zum Ziel aufbereitet worden. Des Weiteren wurden diesmal die Differenzen der horizontalen und vertikalen Joystick-Auslenkung zwischen MWB 1 und 2 aus den Summen der horizontalen und vertikalen Joystick-Auslenkungen relativiert an der Anzahl der Messzeitpunkte gewonnen.

##### **3.2.2.3.2 Analyse der objektiven Tracking-Daten**

Die Auswertung der aufbereiteten Rohdaten erfolgte mittels MS Office Excel 2007, SPSS 19 und GPower 3.1. Die deskriptiven Kennwerte (M, SD, H, %) zur Aufklärung der Forschungsfragestellung 2 sind dem Anhang A.3 zu entnehmen.

Da die Hypothesen keine Unterschiede zwischen den beiden Stufen des Gruppierungsfaktors „Vorselektion“ hinsichtlich der aVn (Differenz bei der horizontalen und vertikalen Joystick-Auslenkung zwischen den MWBn, Fehleranzahl sowie benötigte Zeit beim Tracking)



postulieren, wurde für die inferenzstatistischen Absicherungen der Ergebnisse das Alpha-Fehler-Niveau a priori auf 20 % angehoben<sup>17</sup>.

Aufgrund der kleinen Stichproben von je sechs MWB-Paaren in den beiden Voruntersuchungen, können deren vergleichende Analysen nur bei sehr großen Effektgrößen ausreichend statistische Power besitzen. Um diese Teststärken zu erhöhen, wurden für die statistischen Vergleiche der beiden Voruntersuchungen zweifaktorielle univariate Varianzanalysen<sup>18</sup> mit dem Messwiederholungsfaktor „Durchgang“<sup>19</sup> ausgeführt.<sup>20</sup> Dabei wurden Ausreißer nicht aus den statistischen Analysen ausgeschlossen, denn die Prüfung der Rohdaten ergab keine Messfehler. Zudem kann aufgrund der kleinen Stichproben je Voruntersuchung nicht eindeutig geklärt werden, ob es sich tatsächlich um Ausreißer handelt. Ferner sollten diese Stichproben nicht unnötig dezimiert werden.

### 3.2.2.3.2.1 Ergebnisse der Analysen zur Forschungsfragestellung 2

Nachfolgend werden, aus Platzgründen, nur die hypothesenrelevanten Ergebnisse der Varianzanalysen präsentiert.

#### 2.1 Forschungshypothesen bezüglich $\alpha V$ „Konflikt zwischen den MWBn beim Objekt-Tracking“

Die *Hypothese F.2.1.1* postuliert zwischen den beiden Gabelungsvoruntersuchungen eine Gleichheit hinsichtlich der Differenz der horizontalen Joystick-Auslenkung der beiden MWB beim Tracking. Für den Gruppierungsfaktor „Vorselektion“ (2fach gestuft: 1. Gabelung mit vorselektierten MWBn [Gs], 2. Gabelung mit nicht vorselektierten MWBn [Gu]) zeigt sich

---

<sup>17</sup> Dadurch wird die Wahrscheinlichkeit einer Fehlentscheidung zugunsten einer solchen Gleichheitshypothese (auch als Nullhypothese bekannt) minimiert (Bortz, 1999).

<sup>18</sup> Es existiert kein äquivalentes non-parametrisches Verfahren zur mehrfaktoriellen Varianzanalyse. Nach Bortz (1999) und Zöfel (2003) handelt es sich jedoch bei der Varianzanalyse um ein ziemlich robustes Verfahren gegenüber den Verletzungen von deren Voraussetzungen (Normalverteilung und Varianzhomogenität). Ein nicht signifikantes Ergebnis unter Verletzung der Voraussetzungen würde sich auch bei gegebener Normalverteilung und Varianzhomogenität zeigen (Zöfel, 2003).

Für die Überprüfung der Normalverteilung kamen sowohl der Kolmogorov-Smirnov-Test mit Lilliefors-Schranken als auch der Shapiro-Wilk-Test zum Einsatz. Der Shapiro-Wilk-Test weist eine hohe Teststärke bei  $N < 50$  auf (Seier, 2002). Die Normalverteilung einer Variable wurde nur angenommen, wenn diese durch beide Verfahren mit einem nicht signifikanten Ergebnis ( $p > .20$ ) statistisch bestätigt werden konnte. Die Varianzhomogenität wurde mittels des Levene-Tests überprüft, ebenfalls mit einem Alpha-Fehler-Niveau von 20 %.

<sup>19</sup> In den vier Durchgängen, welche die Stufen des Faktors „Durchgang“ bilden, kamen die vier verschiedenen Gabelungsvarianten zum Einsatz. Dabei wurden vier verschiedene Reihenfolgen verwendet (s. Abschnitt 3.2.1.2.4). Aufgrund der kleinen Stichproben konnten jedoch nicht alle Reihenfolgen mit der gleichen Anzahl an MWB-Paaren besetzt werden. Aus diesem Grund sieht die Autorin von einer Betrachtung der Haupt- und Interaktionseffekte mit dem Faktor „Durchgang“ ab.

<sup>20</sup> Durch die Aufnahme eines weiteren Faktors in eine Varianzanalyse kann die Teststärke für die statistische Überprüfung des Effekts eines Faktors erhöht werden. Dafür muss jedoch die aufgeklärte Variation der abhängigen Variablen durch den zusätzlich aufgenommenen Faktor den jeweiligen Verlust an Freiheitsgraden übersteigen. Zudem ermöglicht jede weitere Stufe des zusätzlichen Faktors die Erhöhung der statistischen Power eines Tests für den überprüften Effekt eines Faktors. (Sedlmeier & Renkewitz, 2008)

kein signifikanter Einfluss auf die Differenz der horizontalen Joystick-Auslenkung der beiden MWB beim Tracking ( $M_{Gs} = 107.52$ ,  $SD_{Gs} = 183.67$ ,  $M_{Gu} = 150.89$ ,  $SD_{Gu} = 218.19$ ,  $F[1, 10] = 0.144$ ,  $p = .712$ ,  $\eta^2_p = .014$ ,  $1 - \beta = .25$ ).

Die **Hypothese F.2.1.2** postuliert zwischen den beiden Gabelungsvoruntersuchungen eine Gleichheit hinsichtlich der Differenz der vertikalen Joystick-Auslenkung der beiden MWB beim Tracking. Für den Gruppierungsfaktor „Vorselektion“ (2fach gestuft: 1. Gabelung mit vorselektierten MWBn [Gs], 2. Gabelung mit nicht vorselektierten MWBn [Gu]) zeigt sich kein signifikanter Einfluss auf die Differenz der vertikalen Joystick-Auslenkung der beiden MWB beim Tracking ( $M_{Gs} = 376.28$ ,  $SD_{Gs} = 483.68$ ,  $M_{Gu} = 550.81$ ,  $SD_{Gu} = 470.12$ ,  $F[1, 10] = 0.590$ ,  $p = .460$ ,  $\eta^2_p = .056$ ,  $1 - \beta = .41$ ).

## **2.2 Forschungshypothese bezüglich aV „Genauigkeit beim Objekt-Tracking“**

Die **Hypothese F.2.2** postuliert zwischen den beiden Gabelungsvoruntersuchungen eine Gleichheit hinsichtlich der Anzahl der Fehler beim Tracking. Für den Gruppierungsfaktor „Vorselektion“ (2fach gestuft: 1. Gabelung mit vorselektierten MWBn [Gs], 2. Gabelung mit nicht vorselektierten MWBn [Gu]) zeigt sich kein signifikanter Einfluss auf die Anzahl der Fehler beim Tracking ( $M_{Gs} = 2346.83$ ,  $SD_{Gs} = 851.44$ ,  $M_{Gu} = 2735.50$ ,  $SD_{Gu} = 936.10$ ,  $F[1, 10] = 0.604$ ,  $p = .455$ ,  $\eta^2_p = .057$ ,  $1 - \beta = .41$ ).

## **2.3 Forschungshypothese bezüglich aV „Geschwindigkeit beim Objekt-Tracking“**

Die **Hypothese F.2.3** postuliert zwischen den beiden Gabelungsvoruntersuchungen eine Gleichheit hinsichtlich der benötigten Zeit zum Tracking. Für den Gruppierungsfaktor „Vorselektion“ (2fach gestuft: 1. Gabelung mit vorselektierten MWBn [Gs], 2. Gabelung mit nicht vorselektierten MWBn [Gu]) zeigt sich kein signifikanter Einfluss auf die benötigte Zeit zum Tracking ( $M_{Gs} = 184433.50$  ms,  $SD_{Gs} = 36189.93$ ,  $M_{Gu} = 162585.59$  ms,  $SD_{Gu} = 50042.59$ ,  $F[1, 10] = 0.954$ ,  $p = .352$ ,  $\eta^2_p = .087$ ,  $1 - \beta = .51$ ).

### **3.2.2.4 Interpretation und Diskussion der Ergebnisse zur Forschungsfragestellung 2**

Zwischen den Voruntersuchungen „Gabelung mit und ohne Vorselektion der MWB nach deren unspezifischen Strategie“ liegen hinsichtlich der aVn „Differenz der horizontalen und vertikalen Joystick-Auslenkung der MWB“, „Fehleranzahl“ sowie „benötigte Zeit“ beim Tracking keine signifikanten Unterschiede vor. Deshalb können die **Hypothesen F.2.1.1**,

---

<sup>21</sup> Konventionen für die Effektgröße  $\eta^2_p$  nach Cohen (1988, zitiert nach Sedlmeier & Renkewitz, 2008):  $\eta^2_p = .01$  ist ein kleiner,  $\eta^2_p = .06$  ist ein mittlerer,  $\eta^2_p = .14$  ist ein großer Effekt.

*F.2.1.2, F.2.2 und F.2.3 nicht verworfen* werden. Die statistischen Überprüfungen weisen jedoch lediglich kleine bis mittlere Teststärken auf. Für sichere statistische Überprüfungen sind somit größere Stichproben notwendig. Die Unterschiede zwischen den Stichproben der beiden Voruntersuchungen hinsichtlich der aVn sind augenscheinlich groß, allerdings liegen ebenso hohe Standardabweichungen vor. Mit Ausnahme bei der Differenz der horizontalen Joystick-Auslenkung der beiden MWB entsprechen diese Unterschiede mittleren Effektgrößen, welche von der Autorin als praktisch bedeutsam erachtet werden. Dabei absolvierte die Stichprobe der nicht vorselektierten MWB die Tracking-Durchgänge durchschnittlich mit einer höheren Differenz bei der horizontalen und vertikalen Joystick-Auslenkung der beiden MWB, mit einer höheren Fehlerrate sowie in kürzerer Zeit. Dieses Ergebnismuster ist für den Einsatz von Operateuren und Automaten geeigneter, als das der Stichprobe der vorselektierten MWB. Durch eine unangemessene Operationalisierung der unspezifischen Strategie sollte das Ergebnis nicht verursacht sein. Zum einen handelt es sich beim FESE-R (Nachtwei, 2008b) um ein empirisch validiertes Messinstrument (s. Anhang A.2.1), zum anderen zeigte sich in der signifikant höheren Nicht-Erscheinungsrate der geschwindigkeitsorientierten Personen bei den Voruntersuchungen (s. Abschnitt 3.2.1.2.6.1), dass die durch den FESE-R (Nachtwei, 2008b) erfasste unspezifische Strategie prinzipiell das Verhalten einer Person beeinflussen kann. Scheinbar wurde die Wirkung der unterschiedlichen unspezifischen Strategie der beiden MWB beim kooperativen Tracking, wie im 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) postuliert, vom Versuchssetting überlagert. Eventuell wirken andere Faktoren, wie die spezifische Strategie, stärker auf das Objektregelungsverhalten der MWB ein. Ebenso kann die kooperativ zu bewältigende Aufgabe selbst dazu geführt haben, dass die MWB eher miteinander kooperierten und dafür ihr Objektregelungsverhalten aneinander, unabhängig von ihrer unspezifischen Strategie, anpassten. Zudem können andere Personenmerkmale, wie die Motivation der MWB, das Ergebnis bedingt haben.

Aufgrund der gefundenen Ergebnisse und der aussichtsarmen Situation unter angemessenem Zeit- und Kostenaufwand eine ausreichende Anzahl an VPn für die noch anstehenden Untersuchungen im ATEO-Projekt zu akquirieren, welche die notwendigen Extremausprägungen hinsichtlich der unspezifischen Strategie aufweisen, wurde entschieden in nachfolgenden Untersuchungen mit SAM auf die Vorselektion der VPn nach deren unspezifischen Strategie zu verzichten. Das Vorgehen schien geeignet, um zukünftig mit SAM versuchsökonomische Untersuchungen durchführen zu können, durch reduzierten Aufwand bei der VPn-Akquise, ohne dabei einen Verlust der durch einen Operateur oder Entwickler wahrnehmbaren Komplexität von SAM zu erleiden.

### 3.2.2.5 Fazit, Kritik und Ausblick

Die Stichproben der einzelnen Voruntersuchungen waren zu klein, um einige Hypothesen statistisch sicher überprüfen zu können. Dennoch erweckten die gefundenen Ergebnisse insgesamt den Anschein, dass die Weiterentwicklung von SAM auf dem richtigen Weg zur Erreichung einer wahrnehmbaren Komplexität war, welche den Einsatz von Entwicklern und Operateuren rechtfertigt. Zudem erwiesen sich der Untersuchungsaufbau, der -ablauf, die -materialien sowie der zeitliche Rahmen als angemessen. Ferner konnten aus den Ergebnissen der Voruntersuchungen einige Optimierungen von SAM abgeleitet werden, welche in der weiterentwickelten Version, SAM 2.0, umgesetzt wurden. So erkannte bspw. in den vier Voruntersuchungen bei zwölf MWB-Paaren (50 % aller Paare) einer der MWB die Unterschiedlichkeit der instruierten spezifischen Strategien für das Objekt-Tracking (s. Anhang A.3.3). Dies könnte durch den Medienbruch zwischen der Instruktion der jeweiligen spezifischen Strategie per Papier und der digitalen Präsentation aller anderen Instruktionen verursacht sein. Die Aufmerksamkeit der MWB wurde dadurch vermutlich zu sehr auf die Briefumschläge und die darin enthaltende „Teamaufgabe“ gelenkt. So sagte eine VP: *„Ja, da Briefumschlag selbst öffnen, muss andere eine andere Instruktion haben“*. Deshalb sind in SAM 2.0 die Auswahl der vermeintlichen „Teamaufgabe“ sowie die Instruktion der beiden spezifischen Strategie digitalisiert und in den Präsentationsablauf der restlichen Instruktionen integriert (s. Anhang B.1.1). Ebenso schien das beobachtbare Solo-Tracking des jeweils anderen MWBs einen Anteil daran zu haben, dass die Instruktion unterschiedlicher spezifischer Strategien aufgedeckt wurde. Eine VP gab an: *„Erst angenommen die Gleiche, aber beim Zuschauen des Fahrens stutzig geworden und gedacht gegensätzliche Instruktion“*. Aus diesem Grund werden in SAM 2.0 nur noch die Test- und Trainingsdurchgänge als Solo-Tracking durchgeführt. Diese werden weiterhin ohne Anwesenheit des anderen MWBs absolviert. Die übrigen Durchgänge müssen jedoch kooperativ bewältigt werden. Des Weiteren könnte das Verbot über die „Teamaufgabe“ zu sprechen, damit ein Beobachter ohne deren Kenntnis das non-verbale Verhalten der MWB per Video auswerten kann, für die MWB verräterisch gewesen sein. Deshalb existiert in SAM 2.0 ein generelles Redeverbot. Zum einen soll dadurch die Aufmerksamkeit der MWB nicht versehentlich auf die „Teamaufgabe“ gelenkt werden. Zum anderen stellt die explizite, aber auch implizite Kommunikation in einem Team eine wichtige Voraussetzung dar, um die Aktivitäten zur Erreichung eines gemeinsamen Ziels zu koordinieren (Skjerve & Skraaning Jr., 2004). Somit sollte ohne eine Möglichkeit zur Kommunikation die Kooperation zwischen den MWBn noch weiter erschwert sein.

Aus der Analyse der erhobenen Kontrollvariablen (Details s. Anhang A.3.3) ergab sich, dass die MWB zwischen sich einen Leistungsunterschied<sup>22</sup> entsprechend deren Strategien wahrgenommen hatten. Diese subjektive Leistungsdifferenz hatte jedoch keinen Einfluss auf deren Zufriedenheit<sup>23</sup>, Stimmung<sup>24</sup> und Sympathie<sup>25</sup> füreinander. Ein solcher Einfluss würde jedoch das Konfliktpotential zwischen den MWBn erhöhen. Eine Ursache könnte sein, dass die MWB nicht ausreichend motiviert waren, um ein gutes Tracking-Ergebnis zu erreichen. Um die extrinsische Motivation der MWB weiter zu erhöhen, wird in der Instruktion von SAM 2.0 mehr Augenmerk auf die Vorstellung der leistungsabhängigen Entlohnung sowie des Gewinns gelegt. Des Weiteren wird im Rahmen der Vorstellung des kooperativen Trackings stärker betont, dass ein sehr gutes Tracking-Ergebnis nur gemeinsam erreicht werden kann. Dabei wird der Teamgedanke stärker in den Vordergrund gestellt.

Die gefundenen Ergebnisse, sowohl hinsichtlich der objektiven Tracking-Daten als auch der subjektiven Kontrollvariablen<sup>26</sup>, erwecken den Eindruck, dass die Vorselektion der VPn hinsichtlich deren unspezifischen Strategie, erhoben durch den FESE-R (Nachtwei, 2008b), im Vergleich zur spezifischen Strategie keinen bedeutsamen Mehrwert besitzt für die Schaffung eines Konflikts zwischen den MWBn. Deshalb wird zukünftig auf die Vorselektion der MWB nach deren unspezifischen Strategie verzichtet.

Die Länge der in SAM 1.5 verwendeten Pfade für das kooperative Tracking würde von den Operateuren und Automaten im Durchschnitt einen Einsatz von drei Minuten pro Tracking-Durchgang erfordern. Ein sehr kurzer Zeitraum für die Operateure und Automaten, um sich auf die komplexen, dynamischen Prozesse von SAM einzustellen. Deshalb ist zusätzlich in SAM 2.0 ein Tracking-Durchgang mit einem ca. dreimal so langen Tracking-Pfad integriert.

---

<sup>22</sup>  $M_{\text{WoraufbeimTrackinggeachtetMWB1selbst}} = 11.91$ ,  $SD_{\text{WoraufbeimTrackinggeachtetMWB1selbst}} = 20.97$ ,  $M_{\text{WoraufbeimTrackinggeachtetMWB1andere}} = -11.63$ ,  $SD_{\text{WoraufbeimTrackinggeachtetMWB1andere}} = 21.46$ ,  $F[1, 20] = 9.660$ ,  $p_{\text{zweiseitig}} = .006$ ,  $\eta^2_p = .326$ ,  $1 - \beta = .84$ ,  $p_{\text{einseitig}} = .003$  und  $M_{\text{WoraufbeimTrackinggeachtetMWB2selbst}} = -26.79$ ,  $M_{\text{WoraufbeimTrackinggeachtetMWB2andere}} = -1.21$ ,  $F[1, 20] = 28.233$ ,  $p_{\text{zweiseitig}} < .001$ ,  $\eta^2_p = .585$ ,  $1 - \beta > .999$ ,  $p_{\text{einseitig}} < .001$

<sup>23</sup> 83.30 % der MWB waren eher bis sehr zufrieden mit der Teamleistung nach den kooperativen Tracking-Durchgängen.

<sup>24</sup>  $M_{\text{StimmungMWBPrä}} = 5.42$ ,  $SD_{\text{StimmungMWBPrä}} = 1.15$ ,  $M_{\text{StimmungMWBPst}} = 5.67$ ,  $SD_{\text{StimmungMWBPst}} = 0.88$ ,  $F[1, 20] = 1.127$ ,  $p_{\text{zweiseitig}} = .301$ ,  $\eta^2_p = .053$ ,  $1 - \beta = .24$ ,  $p_{\text{einseitig}} = .151$

<sup>25</sup>  $M_{\text{SympathieMWBPrä}} = 5.61$ ,  $SD_{\text{SympathieMWBPrä}} = 0.90$ ,  $M_{\text{SympathieMWBPst}} = 5.92$ ,  $SD_{\text{SympathieMWBPst}} = 0.92$ ,  $F[1, 20] = 5.000$ ,  $p_{\text{zweiseitig}} = .037$ ,  $\eta^2_p = .200$ ,  $1 - \beta = .57$ ,  $p_{\text{einseitig}} = .019$ . Die Stichproben weisen keine Varianzhomogenität auf, deshalb wurde die Signifikanzschranke auf .01 gesetzt (s. Bühl & Zöfel, 2005).

<sup>26</sup> Es konnten keine signifikanten Unterschiede zwischen den vier Voruntersuchungen von SAM 1.5 hinsichtlich der analysierten Variablen gefunden werden.

### **3.2.3 Hauptuntersuchung: Vergleich SAM 1.0 vs. 2.0**

Den vier Voruntersuchungen folgte eine Hauptuntersuchung. Diese diente einem umfangreichen Manipulation-Check. Geprüft werden sollte, ob die Erweiterungen der Versuchsumgebung den Bedarf an einem Operateur oder Automaten zur Prozessführung von SAM 2.0 erhöht haben im Vergleich zu SAM 1.0. Dafür wurde die objektive Tracking-Leistung der MWB in SAM 2.0 mit der in SAM 1.0 experimentell verglichen. Die analysierte Tracking-Leistung der MWB in SAM 2.0 wurde in einer Untersuchung von der Autorin erhoben. Die relevanten Tracking-Verhaltensdaten der MWB in SAM 1.0 wurden bereits in einer Untersuchung von Gross erfasst.

Ebenso wurde ermittelt, ob längere Tracking-Pfade zu einer Erhöhung der durch Entwickler oder einen Operateur wahrnehmbaren Komplexität von SAM führen. Dafür wurden mit Hilfe eines kurzen und eines langen Pfads in SAM 2.0 die objektive Tracking-Leistung der MWB, deren subjektive Beanspruchung und Komplexitätswahrnehmung der Tracking-Aufgabe verglichen. In der Untersuchung von SAM 2.0 sind die Operationalisierungen der aV „Geschwindigkeit“ und der Kontrollvariablen, der Untersuchungsort sowie die Erhebungsinstrumente identisch mit denen der vier Voruntersuchungen (s. Abschnitt 3.2.2.2.2) sowie der Untersuchung von SAM 1.0. Deshalb werden zur Vermeidung von Redundanzen nachfolgend nur die Unterschiede erwähnt.

#### **3.2.3.1 Fragestellungen und Hypothesen**

In diesem Abschnitt sollen die Herleitung der Forschungs- und Kontrollfragestellungen vorgestellt werden, die dem Vergleich SAM 1.0 und 2.0 zugrunde liegen. Hierbei handelt es sich um methodische Fragestellungen, welche sich auf die Weiterentwicklung der Versuchsumgebung beziehen. Diese stellen keine inhaltlichen Forschungsfragestellungen dieser Dissertation oder des ATEO-Projekts dar. Zusätzlich werden in diesem Abschnitt die aus den Fragestellungen abgeleiteten Hypothesen präsentiert. Bei diesen handelt es sich um Gleichheits- oder gerichtete, unspezifische Unterschieds- und Veränderungshypothesen.

##### **3.2.3.1.1 Herleitung der dritten Forschungsfragestellung**

Im Bemühen die wahrnehmbare Komplexität der Versuchsumgebung SAM zu erhöhen, wurde diese von der Version 1.0 zu 2.0 inhaltlich und technisch weiterentwickelt. Dabei wurde der Fokus auf die Erhöhung des Konflikts zwischen den MWBn gerichtet. Dafür erhalten die beiden MWB eine konträre spezifische Strategie für das Tracking, welche ihnen jedoch als eine gemeinsame „Teamaufgabe“ vorgestellt wird (s. Abschnitt 3.1.2). Ferner wird den MWBn nun diese vermeintliche „Teamaufgabe“ zusammen mit allen anderen Instruktionen von SAM

digital präsentiert (s. Abschnitt 3.2.2.5). Dadurch soll verhindert werden, dass die MWB ihrer spezifischen Strategien eine besondere Aufmerksamkeit schenken, welche zur Aufdeckung des Unterschieds zur spezifischen Strategie des anderen MWBs führen kann. Ebenso soll die Leistungsmotivation der MWB durch ein permanentes Leistungsfeedback über die Geschwindigkeit und die Genauigkeit während des Trackings, eine permanent visualisierte leistungsbabhängige Entlohnung sowie die Chance auf einen Gewinn von 100 Euro für das leistungstärkste MWB-Team erhöht werden (s. Abschnitt 3.1.1). Überdies sind die Gabelungen so verändert, dass die MWB in der Lage sind, sich für eine der Pfadoptionen aufgrund deren Gestaltung zu entscheiden (s. Abschnitt 3.1.3.1). Wobei eine Pfadoption durch den genauigkeitsorientierten und die andere durch den geschwindigkeitsorientierten MWB präferiert werden sollte. Zudem sind zwei statische und ein dynamisches Hindernis in den Tracking-Pfad integriert (s. Abschnitt 3.1.3.2). Diese sollten den MWBn ebenfalls eine Entscheidung zwischen einem schnellen oder genauen Regeln des Tracking-Objekts abverlangen.

Aufgrund dieser Veränderungen in SAM 2.0 sollten die beiden gegensätzlichen spezifischen Strategien der beiden MWB beim Tracking mehr zum Tragen kommen als bei SAM 1.0. Dabei sollten sich die unterschiedlichen spezifischen Strategien im unterschiedlichen Objektregelungsverhalten der beiden MWB widerspiegeln. Die Wahrnehmung dieser unterschiedlichen Tracking-Verhaltensweisen, trotz einer vermeintlich gemeinsamen „Teamaufgabe“, sollte einen Konflikt zwischen den MWBn verursachen. Ein Konflikt führt zu nicht zielorientierten Verhaltensweisen der MWB (s. Jehn, Chadwick & Thatcher, zitiert nach Nachtwei, 2008a). Dadurch wird die Varianz der Verhaltensweisen der MWB erhöht und damit ebenso die wahrnehmbare Komplexität von SAM 2.0 im Vergleich zu SAM 1.0. Es ergab sich somit folgende Fragestellung:

**Forschungsfragestellung 3:**

Erhöhen die vorgenommenen Veränderungen in SAM 2.0 dessen spezifisch auf Konflikte der MWB bezogene Komplexität im Vergleich zu SAM 1.0.?

Die wahrnehmbare erhöhte Komplexität von SAM 2.0, aufgrund der stärkeren Wirkung der beiden konträren spezifischen Strategien und der sich daraus ergebenden konfliktbedingten vielfältigen Verhaltensweisen der MWB, sollte sich in Summe in einer schnelleren, aber gleichzeitig ungenaueren Objektregelung durch die MWB zeigen. Durch eine solche Tracking-Leistung der MWB sollte die Komplexität von SAM durch die Entwickler oder den Operateur als höher wahrgenommen werden. Damit einher sollte eine höhere Anforderung an den Operateur

oder die Automaten bei der Prozessführung von SAM 2.0 gehen. Daraus ergaben sich die nachfolgenden zwei Hypothesen zur Aufklärung der Forschungsfragestellung 3:

### **3.1 Forschungshypothese bezüglich $\alpha V$ „Genauigkeit beim Objekt-Tracking“**

*F.3.1: Es existiert ein Unterschied bei der Anzahl an Fehlern beim Tracking zwischen den beiden Untersuchungen SAM 1.0 und 2.0. Unter Einsatz von SAM 2.0 verursachen die MWB mehr Fehler beim Tracking.*

### **3.2 Forschungshypothese bezüglich $\alpha V$ „Geschwindigkeit beim Objekt-Tracking“**

*F.3.2: Es existiert ein Unterschied bei der benötigten Zeit zum Tracking zwischen den beiden Untersuchungen SAM 1.0 und 2.0. Unter Einsatz von SAM 2.0 benötigen die MWB weniger Zeit zum Tracking.*

#### **3.2.3.1.2 Herleitung der vierten Forschungsfragestellung**

In SAM 2.0 existieren ein kurzer und ein dreimal so langer Tracking-Pfad für das kooperative Tracking. Diese bestehen aus verschiedenen Elementen (Geraden, Kurven, Gabelungen, Hindernisse), um Varianz im Verlauf zu erreichen. Trotzdem konnten Wiederholungen der Elemente sowie deren Kombinationen nicht vermieden werden. Diese Wiederholungen erleichtern den MWBn das Erlernen der Objektregelung (s. Russell & Sternad, 2001, zitiert nach Gross, 2008). Bei dem dreimal so langen Tracking-Pfad könnte dieser Lerneffekt aufgrund der größeren Anzahl an Wiederholungen größer ausfallen. Dadurch würden die MWB über den Tracking-Pfad hinweg weniger beansprucht werden. Deren Tracking-Leistung könnte ansteigen. Gleichzeitig würden sich die wahrnehmbare Komplexität von SAM und damit der Einsatzbedarf von Operateuren und Automaten verringern.

Ebenso könnte der dreimal so lange Tracking-Pfad die Anforderungen an die kognitive und motorische Leistung der MWB erhöhen. Diese erhöhte Belastung der MWB könnte eine subjektive Beanspruchung und Ermüdung der MWB zur Folge haben. Dies könnte kurzfristig zu Konflikten zwischen diesen sowie der Minderung von deren Leistung führen (s. Sarodnick & Brau, 2006). Dadurch wiederum würde die wahrnehmbare Komplexität von SAM ansteigen. Andererseits könnte eine höhere subjektive Beanspruchung der MWB ebenso deren Motivation vermindern, die Tracking-Aufgabe bzw. die spezifische Strategie auszuführen. Dadurch würde sich die Komplexität von SAM ebenfalls verringern. Zudem könnte dies in Untersuchungen mit Einsatz eines Operateurs oder von Automaten zur Regelung des Prozesses von SAM dazu führen, dass die MWB deren Unterstützung nicht annehmen. Aus diesen Gründen muss für die Versuchsumgebung die Pfadlänge identifiziert werden, welche eine subjektive Beanspruchung



der MWB zur Folge hat, mit der eine Komplexitätserhöhung von SAM einhergeht. Aus diesem Anliegen leitete sich folgende Fragestellung ab:

**Forschungsfragestellung 4:**

Bei welcher Pfadlänge weist SAM eine höhere Komplexität auf?

Für die Aufklärung dieser Fragestellung wird ein exploratives Vorgehen gewählt. Es liegt keine Annahme darüber vor, welche der beiden Pfadlängen geeigneter ist, die wahrnehmbare Komplexität von SAM zu erhöhen.

**3.2.3.1.3 Herleitung der ersten Kontrollfragestellung**

Die Umsetzung der beiden konträren spezifischen Strategien der beiden MWB beim Tracking ist die Basis für die Wahrnehmung einer Leistungsdivergenz zwischen ihnen. Der objektive Leistungsunterschied sollte durch die Steigerung der Leistungsmotivation verstärkt werden. Hierfür wurden ein permanentes Leistungsfeedback über die Geschwindigkeit und Genauigkeit während des Trackings, eine leistungsabhängige Entlohnung der MWB und die Aussicht auf einen 100 Euro-Gewinn für das MWB-Team mit der besten Tracking-Leistung verwendet (s. Abschnitt 3.1.1). Ebenfalls sollte durch das permanente Leistungsfeedback sowie die permanent sichtbaren leistungsabhängigen Entlohnungsstufen diese Motivation über die gesamte Untersuchung aufrecht erhalten bleiben. Daher stellte sich folgende Kontrollfrage:

**Kontrollfragestellung 1:**

Motivieren die geschaffenen Rahmenbedingungen der Versuchsumgebung SAM 2.0 die MWB dazu die spezifische Strategie im Tracking-Verhalten umzusetzen?

Nachfolgend werden die aus der Kontrollfragestellung abgeleiteten Hypothesen vorgestellt:

**1.1 Kontrollhypothese bezüglich Instruktionsverständlichkeit**

Die Grundvoraussetzung für eine Befolgung der instruierten spezifischen Strategie ist deren Verständlichkeit. Eine kurze, prägnante Formulierung sowie Visualisierung dieser (s. Anhang A.1.1) sollen dafür sorgen, dass diese verständlich ist. Daraus ergab sich die Hypothese:

*K.1.1: Es existiert kein Unterschied in der subjektiven Instruktionsverständlichkeit zwischen MWB 1 und 2. Beide MWB beurteilen die Instruktion der spezifischen Strategie als verständlich.*

### **1.2 Kontrollhypothese bezüglich Instruktionsbefolgung**

Die Motivation, die instruierte spezifische Strategie umzusetzen, spiegelt sich in deren Befolgung wider. Somit lautete die Hypothese:

*K.1.2: Es existiert kein Unterschied in der subjektiven Instruktionsbefolgung zwischen MWB 1 und 2. Beide MWB geben an, die Instruktion der spezifischen Strategie befolgt zu haben.*

### **1.3 Kontrollhypothese bezüglich der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen**

Eine hohe Motivation die spezifische Strategie umzusetzen, sollte sich zudem darin zeigen, dass es den MWBn wichtig ist, ein gutes Tracking-Ergebnis zu erreichen. Daraus ergab sich die Hypothese:

*K.1.3: Es existiert kein Unterschied bei der Selbstbeurteilung der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen zwischen MWB 1 und 2. Beiden MWBn ist es wichtig ein gutes Ergebnis zu erzielen.*

### **1.4 Kontrollhypothesen bezüglich worauf beim Tracking geachtet wurde**

Die Motivation eines Menschen schlägt sich in einem zweckorientierten Handeln nieder (Müsseler & Prinz, 2002). Somit sollten die MWB entsprechend deren spezifischen Strategien eine Präferenz aufweisen, schnell oder genau zu regeln. Daher lautete die Hypothese:

*K.1.4.1: Es existiert ein Unterschied im Selbsturteil von MWB 1 und 2, worauf beim Tracking mehr geachtet wurde: Geschwindigkeit vs. Genauigkeit. MWB 1 hat mehr auf die Geschwindigkeit geachtet. Dagegen hat MWB 2 mehr auf die Genauigkeit geachtet.*

Überdies bewirkt die menschliche Motivation ein ausdauerndes Handeln (Müsseler & Prinz, 2002). Somit sollte diese Präferenz ebenso zeitlich sehr stabil sein. Wenn jedoch die Motivation der MWB, die jeweilige spezifische Strategie umzusetzen, über den Untersuchungsverlauf abnimmt, könnte dies wiederum eine Verringerung des Konflikts zwischen diesen bedeuten. Daher wurde folgende Hypothese aufgestellt:

*K.1.4.2: Es existieren keine Unterschiede zwischen den Tracking-Durchgängen im Selbsturteil der MWB, worauf beim Tracking mehr geachtet wurde.*

#### **3.2.3.1.4 Herleitung der zweiten Kontrollfragestellung**

Nach dem 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) führen kooperierende Personen einen Vergleich der Einzelleistungen durch. Dadurch sollte die, durch die konträren spezifischen Strategien der beiden MWB verursachte unterschiedliche Objektregelung der MWB durch diese wahrgenommen werden. Dieser subjektive Leistungsunterschied zwischen den MWBn sollte einen Konflikt zwischen diesen verursachen. Aus diesen Bemühungen leitete

sich folgende Kontrollfragestellung ab:

**Kontrollfragestellung 2:**

Verursacht die subjektive Leistungsdifferenz der MWB beim Tracking einen Konflikt zwischen ihnen?

Für diese Kontrollfragestellung ließen sich nachfolgende Hypothesen formulieren:

**2.1 Kontrollhypothesen bezüglich worauf beim Tracking geachtet wurde**

Damit ein Konflikt zwischen den MWBn entstehen kann, müssen diese erkennen, dass der andere MWB mehr auf das andere Leistungskriterium beim Tracking achtet. Daraus leiteten sich folgende Hypothesen ab:

*K.2.1.1: Bei MWB 1 existiert ein Unterschied bei der Selbstbeurteilung und der Beurteilung des anderen MWBs, worauf beim Tracking geachtet wurde. MWB 1 gibt an, selbst auf die Geschwindigkeit geachtet zu haben. Wohingegen dieser einschätzt, MWB 2 hat eher auf die Genauigkeit geachtet.*

*K.2.1.2: Bei MWB 2 existiert ein Unterschied bei der Selbstbeurteilung und der Beurteilung des anderen MWBs worauf beim Tracking geachtet wurde. MWB 2 gibt an, selbst auf die Genauigkeit geachtet zu haben. Wohingegen dieser einschätzt, MWB 1 hat eher auf die Geschwindigkeit geachtet.*

**2.2 Kontrollhypothese bezüglich der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen**

Die Entstehung eines Konflikts zwischen den MWBn wird gefördert, wenn diese den Eindruck gewinnen, dass dem anderen MWB ein gutes Tracking-Ergebnis weniger wichtig ist als ihnen selbst. Daher ergab sich folgende Hypothese:

*K.2.2: Bei den beiden MWBn existiert ein Unterschied bei der Selbstbeurteilung und der Beurteilung des anderen MWBs, wie wichtig es war, ein gutes Tracking-Ergebnis zu erreichen. MWB 1 und 2 beurteilen die Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen jeweils für sich selbst höher als für den anderen MWB.*

**2.3 Kontrollhypothese bezüglich der Zufriedenheit**

Ein Konflikt zwischen den MWBn sollte zu deren Unzufriedenheit mit der Teamleistung führen (s. Jehn, Chadwick & Thatcher, 2000, zitiert nach Nachtwei, 2008a). Deshalb wurde nachfolgende Hypothese aufgestellt:

*K.2.3: Es existiert kein Unterschied bei der Selbstbeurteilung der Zufriedenheit nach allen kooperativen Tracking-Durchgängen zwischen dem MWB 1 und 2. Die Zufriedenheit fällt bei beiden MWBn eher mittelmäßig bis gering aus.*

## **2.4 Kontrollhypothesen bezüglich der Stimmung**

Ebenso verursachen Konflikte negative Emotionen (Badke-Schaub & Frankenberger, 2004). Somit sollte sich ein Konflikt zwischen den MWBn auf deren Stimmung auswirken. Daher ergaben sich folgende Hypothesen:

*K.2.4.1: Es existiert kein Unterschied bei der Selbstbeurteilung der Stimmung zwischen dem MWB 1 und 2.*

*K.2.4.2: Es existiert ein Unterschied hinsichtlich der Selbstbeurteilung der Stimmung der beiden MWB vor und nach allen kooperativen Tracking-Durchgängen. Die Stimmung der MWB nimmt ab.*

## **2.5 Kontrollhypothese bezüglich Gleichheit der instruierten Tracking-Strategie**

Durch die Annahme der beiden MWB, dieselbe Tracking-Strategie zu verfolgen, sollte der aufgrund des unterschiedlichen Objektregelungsverhaltens entstandene Konflikt gesteigert werden. Daher wurde nachfolgende Hypothese aufgestellt:

*K.2.5: Bei allen MWB-Paaren gehen beide MWB die gesamte Untersuchung davon aus, dieselbe spezifische Strategie für das Tracking instruiert bekommen zu haben.*

## **3.2.3.2 Methoden**

### **3.2.3.2.1 Operationalisierung der Variablen**

Der folgende Abschnitt soll dem Leser dazu dienen, sich einen Überblick über die Operationalisierungen der uVn, der aVn und der Kontrollvariablen zu verschaffen.

#### **3.2.3.2.1.1 Operationalisierung der unabhängigen Variablen**

Nachfolgend werden die uVn vorgestellt, welche für die inferenzstatistische Aufklärung der Forschungs- und Kontrollfragestellungen verwendet wurden.

#### **1. uV „SAM-Version“**

Diese uV bestand aus zwei Stufen, den beiden Untersuchungsgruppen. Diese unterschieden sich in der erlebten SAM-Version:

*Untersuchungsgruppe 1: SAM 1.0*

*Untersuchungsgruppe 2: SAM 2.0*

Es handelte sich somit um einen Between-Faktor bzw. Gruppierungsfaktor. Die Daten der ersten Untersuchungsgruppe wurden in der ersten ATEO-Projektphase von Gross unter Einsatz von SAM 1.0 in einer nicht veröffentlichten Arbeit erhoben und der Autorin zur Verfügung gestellt. Für die zweite Untersuchungsgruppe wurde in der zweiten ATEO-Projektphase

die Untersuchung der ersten Untersuchungsgruppe wiederholt. Jedoch wurden unerfahrene VPn sowie SAM 2.0 eingesetzt.

## **2. uV „Tracking-Pfadlänge“**

Diese uV bestand ebenfalls aus zwei Stufen, den zwei unterschiedlich langen kooperativen Tracking-Pfaden:

*1x solange wie kooperativer Tracking-Pfad SAM 1.5 (kTP)*

*3x solange wie kooperativer Tracking-Pfad SAM 1.5 (lTP)*

Es handelte sich somit um einen Within-Faktor bzw. Messwiederholungsfaktor.

## **3. uV „Durchgang“**

Diese uV bestand aus sechs Stufen, den kooperativen Tracking-Durchgängen mit Gabelungen und Hindernissen:

*kooperativer Tracking-Durchgang 7 (D7)*

*kooperativer Tracking-Durchgang 8 (D8)*

*kooperativer Tracking-Durchgang 10 (D10)*

*kooperativer Tracking-Durchgang 12 (D12)*

*kooperativer Tracking-Durchgang 14 (D14)*

*kooperativer Tracking-Durchgang 16 (D16)*

Es handelte sich somit um einen Within-Faktor bzw. Messwiederholungsfaktor.

## **4. uV „Messzeitpunkt“**

Diese uV bestand ebenfalls aus sechs Stufen, den Messzeitpunkten einiger Kontrollvariablen vor oder nach den messrelevanten kooperativen Tracking-Durchgängen mit Gabelungen und Hindernissen:

*Messzeitpunkt nach dem Probedurchgang (MPrä1)*

*Messzeitpunkt nach dem kooperativen Durchgang 6 (MPrä2)*

*Messzeitpunkt nach dem kooperativen Durchgang 8 (MD8)*

*Messzeitpunkt nach dem kooperativen Durchgang 12 (MD12)*

*Messzeitpunkt nach dem kooperativen Durchgang 16 (MD16)*

*Messzeitpunkt nach allen Durchgängen (MPost)*

Es handelte sich somit um einen Within-Faktor bzw. Messwiederholungsfaktor.

### **5. uV „MWB“**

Diese uV bestand aus zwei Stufen, den beiden MWBn:

*MWB 1*

*MWB 2*

Es handelte sich somit um einen Messwiederholungsfaktor.

### **6. uV „MWB selbst vs. anderer“**

Diese uV bestand ebenfalls aus zwei Stufen, der Selbstbeurteilung und der Beurteilung des jeweils anderen MWBs:

*Selbsturteil eines MWBs (selbst)*

*Fremdurteil eines MWBs über anderen MWB (andere)*

Es handelte sich somit um einen Within-Faktor bzw. Messwiederholungsfaktor.

#### **3.2.3.2.1.2 Operationalisierung der abhängigen Variablen**

Das Ziel der Erweiterungen von SAM war es, deren wahrnehmbare Komplexität durch Minimierung der Kooperation respektive Maximierung der Konflikte zwischen den MWBn zu erhöhen. Kooperation kann durch externe (objektive Kriterien) sowie interne (subjektive Kriterien) Leistung gemessen werden (Hoc, 2001). Aus diesem Grund wurden objektive Tracking-Daten als auch subjektive Beurteilungsdaten nach einem Tracking-Durchgang als aVn verwendet.

#### **Objektive Tracking-Daten**

##### **1. aV „Genauigkeit“**

Die Genauigkeit wurde diesmal durch die aV „Flächenfehler“ operationalisiert. Dafür wurden für die direkt aufeinanderfolgenden Messzeitpunkte die Abweichungen des Objektmittelpunkts von der Mittellinie als Fläche in der Einheit von Pixeln (px) bestimmt. Die Summe der Abweichungsflächen vom Start bis zum Ziel ergibt den Flächenfehler.

In der ersten ATEO-Projektphase wurden verschiedene Fehlermaße verwendet und deren Nützlichkeit geprüft (s. Nachtwei, 2008a). Der Flächenfehler erwies sich als noch präziser zur Darstellung der Tracking-Genauigkeit als der Fehler „NrSensor“. Deshalb diente dieser in der verwendeten Untersuchung „SAM 1.0“ der ersten ATEO-Projektphase als Operationalisierung der Genauigkeit. Weil die Genauigkeit ebenso für den Vergleich der beiden SAM-Versionen als aV verwendet wurde, kam der Flächenfehler ebenfalls für SAM 2.0 zum Einsatz. Die aV „Genauigkeit“ diente ferner der Überprüfung des Messwiederholungsfaktors „Tracking-Pfadlänge“.

## **2. aV „Geschwindigkeit“**

Die Operationalisierung der aV „Geschwindigkeit“ findet sich im Abschnitt 3.2.1.2.1.2. Diese aV kam für den Vergleich der beiden SAM-Versionen sowie der unterschiedlich langen Tracking-Pfade zum Einsatz.

## **Subjektive Fragebogendaten**

### **1. aV „Beanspruchung“**

Die Operationalisierung der aV „Beanspruchung“ erfolgte über die subjektive Einschätzung der MWB, wie anstrengend die Tracking-Aufgabe war (s. Tabelle 10 und Anhang B.3.1). Diese aV diente dem Vergleich der beiden unterschiedlich langen Tracking-Pfade.

### **2. aV „Komplexität“**

Die aV „Komplexität“ wurde ebenfalls über die subjektive Beurteilung der MWB operationalisiert, wie komplex diese die Tracking-Aufgabe empfanden (s. Tabelle 10 und Anhang B.3.1). Diese aV wurde ebenso für den Vergleich der beiden unterschiedlich langen Tracking-Pfade verwendet.

### **3.2.3.2.1.3 Operationalisierung der Kontrollvariablen**

Die Operationalisierung der Kontrollvariablen zur Aufklärung der Kontrollfragestellungen sind der Tabelle 10 zu entnehmen. Zusätzlich wurden in der Untersuchung von SAM 2.0 erneut die Kontrollvariablen mit potentielltem Einfluss auf die aVn aus den Voruntersuchungen erfasst und durch dieselben Items operationalisiert (s. Anhang B.3.1), da das Verhalten der MWB multipel determiniert ist (s. Bargh & Ferguson, 2000). Neben den uVn können weitere Faktoren Einfluss auf die beiden MWB und deren Objektregelungsverhalten haben, wie bspw. die Müdigkeit der MWB beim Tracking. Diese Effekte können zu einer nicht intendierten Varianz bei den aVn führen, welche post hoc durch den Einbezug von Kontrollvariablen in die inferenzstatistischen Analysen aufgeklärt werden kann (Jo, 2002). Dadurch können Ergebnisse eindeutiger interpretiert (Erhöhung der internen Validität) und somit sicherere Aussagen über die Wirkung der uVn getroffen werden (Huber, Kain, & Nachtwei, 2008; Kain & Nachtwei, 2009; C. Meyer, Nachtwei, & Kain, 2009; Nachtwei & Kain, 2010). Aus diesem Grund wurde diesmal auch die unspezifische Strategie als Kontrollvariable erhoben, um deren Einfluss auf die aVn zu prüfen.

In Tabelle 10 sind die im Rahmen dieser Arbeit relevanten Kontrollvariablen der Untersuchung von SAM 2.0 samt Antwortskalen und Messzeitpunkte aufgelistet. Anhang B.3.1 bietet eine Übersicht über alle erhobenen Kontrollvariablen in der Untersuchung von SAM 2.0.

Die konkreten Formulierungen der Items und deren Antwortskalen sind im Anhang A.2.4 nachzulesen. Einige der Kontrollvariablen wurden zu verschiedenen Zeitpunkten erhoben, um Veränderungen über die Zeit zu ermitteln. Bei den Kontrollvariablen wurde darauf geachtet, dass diese ökonomisch messbar sind<sup>27</sup>, indem diese jeweils durch ein Item oder eine Kurzskala<sup>28</sup> operationalisiert wurden. Die MWB beantworteten die Items unabhängig voneinander.

Tabelle 10: Die, für die Datenanalyse, relevanten erhobenen Variablen respektive eingesetzten Fragebogen, die verwendeten Antwortskalen und Messzeitpunkte in der Untersuchung von SAM 2.0

Variable (Fragebogen)	Antwortskala	Prä1	Prä2	nach Durchgang								Post	Ab
				7	8	10	12	14	16	17			
abhängige Variable													
Beanspruchung (RSE <sub>A</sub> )	VAS (0 bis 150)	X		X	X	X	X	X	X	X	X		
Aufgabenkomplexität	VAS (-50 bis 50)			X	X	X	X	X	X	X	X		
Kontrollvariable zur Aufklärung der Kontrollfragestellung 1 und 2													
spezifische Strategie Instruktionsverständnis	VAS (-50 bis 50)											X	
spezifische Strategie Instruktionsbefolgung	VAS (-50 bis 50)											X	
Präferenz schnell vs. genau (selbst)	VAS (-50 bis 50)	X			X		X		X			X	
Präferenz schnell vs. genau (andere)	VAS (-50 bis 50)											X	
Wichtigkeit der Teamleistung (selbst)	VAS (-50 bis 50)											X	
Wichtigkeit der Teamleistung (andere)	VAS (-50 bis 50)											X	
Zufriedenheit mit Teamleistung	VAS (-50 bis 50)											X	
Stimmung	Kunin (7-stufig)		X									X	
Vermutung spezifische Strategie des Anderen	offen												X
Änderung Vermutung spezifische Strategie des Anderen	offen												X
Weitere Kontrollvariablen													
Müdigkeit	VAS (-50 bis 50)		X									X	
Sicherheit im Umgang mit Joystick	VAS (-50 bis 50)		X									X	
unspezifische Strategie (FESE-R von Nachtwei, 2008b)	VL (4- bis 5- stufig)											X	
Anmerkungen. Prä1 = nach Tracking-Durchgang 1 bzw. 3; Prä2 = nach Tracking-Durchgang 6; Post = nach letztem Tracking-Durchgang; Ab = Abschlussbefragung; VAS = visuelle Analogskala; VL = Ratingskala mit verbalem Label für jede Stufe													

<sup>27</sup> Aus dieser Bemühung entstand das von Nachtwei, Huber und der Autorin ins Leben gerufene Projekt „Persönlichkeitsmerkmale als Kontrollvariablen in der Human Factors Forschung“ (PeKoF). Im Rahmen dieses Projekts werden für die Human Factors Forschung relevante Kurzskalen anderen Forschern im Bereich Human Factors kostenlos mit umfangreichem Informationsmaterial bereitgestellt (Detaillierte Informationen s. Anhang A.2.1; Huber et al., 2008; Kain & Nachtwei, 2009).

<sup>28</sup> Kurzskalen stellen eine angemessene Alternative zu langen Fragebogen dar, wenn eine versuchsökonomische Erhebung von Nöten ist (Herzberg & Brähler, 2006).



### 3.2.3.2.2 Erhebungsinstrumente

#### 3.2.3.2.2.1 Rating Scale Effort <sub>ATEO</sub> (RSE<sub>A</sub>)

Für die Erfassung wie beanspruchend eine Tracking-Aufgabe von den MWBN empfunden wurde, wurden von Nachtwei (2011a) und der Autorin Anpassungen an der englischsprachigen Rating Scale Mental Effort (RSME) von Zijstra und van Doorn (1985) vorgenommen (s. Abbildung 8). Bei der RSME-Skala handelt es sich um eine Intervallskala mit den Polen „0“ und „150“ sowie numerischen Ankern in äquidistanten Abständen (Zehnerschritte).

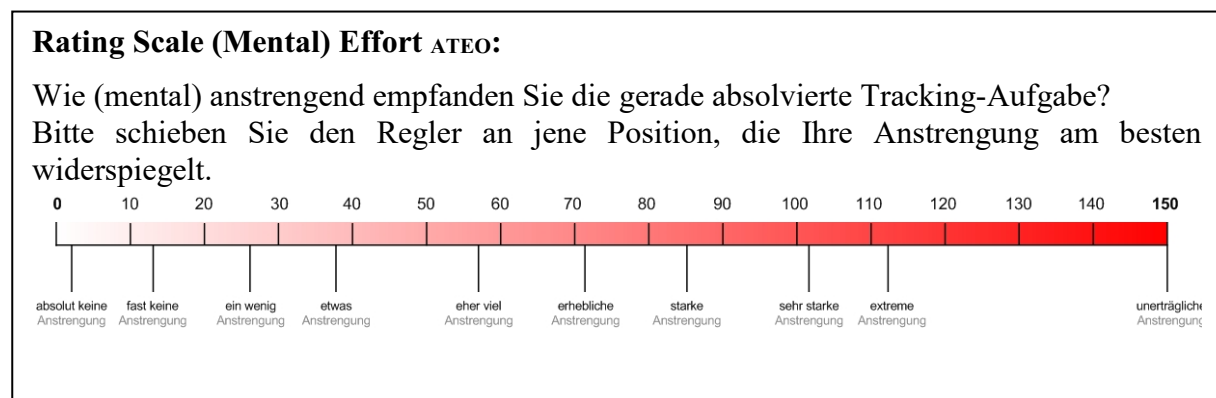


Abbildung 8: Rating Scale Mental Effort <sub>ATEO</sub> und Rating Scale Effort <sub>ATEO</sub>

An der Skala befinden sich zudem in nicht äquidistanten Abständen neun verbale Anker. Für diese Skala existiert neben der englischsprachigen eine niederländische Version (Zijlstra, 1993) sowie eine deutsche Version von Eilers, Nachreiner und Hänecke (1986), die Skala zur Erfassung subjektiv erlebter Anstrengung (SEA). Letztere weicht über die Sprache hinaus hinsichtlich der numerischen und verbalen Anker vom Original ab. Um möglichst nah an der Originalskala zu bleiben, wurden von Nachtwei (2011a) und der Autorin die verbalen Anker aus dem Englischen ins Deutsche übersetzt. Zudem wurde eine horizontale Achse für die Skala verwendet. Im Gegensatz zur Originalskala, welche vertikal orientiert ist. Dadurch sollte der natürliche Lesefluss von links nach rechts unterstützt werden. Des Weiteren wurde ein zehnter verbaler Anker am Ende der Skala bei dem Wert 150 eingeführt, um die Interpretierbarkeit dieses Endpunkts eindeutig zu gestalten. Dies sollte wiederum die Fehlervarianz bei der Urteilsabgabe minimieren. Zusätzlich wurde die auf der Skala abgebildete steigende Intensität der subjektiven Beanspruchung durch eine farbliche Kodierung visuell verstärkt. Dabei beginnt die Skala bei dem Nullwert mit einem neutralen Weiß und wechselt dann in ein immer intensiver werdendes Rot bis zum Endpunkt. Durch zwei verschiedene Textvarianten wurden zwei Skalen-Varianten geschaffen. Zum einen entstand die Version „Rating Scale Mental Effort <sub>ATEO</sub>“ (RSME<sub>A</sub>) weiterhin zur Erfassung der subjektiven mentalen Beanspruchung. Zum

anderen entstand die Version „Rating Scale Effort<sub>ATEO</sub>“ (RSE<sub>A</sub>) zur Erfassung der subjektiven Beanspruchung im Allgemeinen.

Die RSME<sub>A</sub> wurde zur Bestimmung der konvergenten Validität<sup>29</sup> zusammen mit der deutschen ungewichteten Kurzversion des NASA Task Load Index (NASA-TLX, Grubert, n. D.) in der, im zweiten Teil dieser Dissertation beschriebenen, Untersuchung mit 90 Entwicklern eingesetzt. Bei dem NASA-TLX (englisches Original von Hart & Staveland, 1988) handelt es sich um einen in Forschung und Praxis etablierten Fragebogen zur Erhebung der subjektiven Beanspruchung mit den sechs Dimensionen: „geistige Anforderung“, „körperliche Anforderung“, „zeitliche Anforderung“, „Aufgabenausführung“, „Anstrengung“ und „Frustration“. Die verwendete Kurzversion erhebt jede Dimension mit je einem Item (Details s. Anhang G.3.2). Wie erwartet, ergab sich der größte positive Zusammenhang zwischen der **RSME<sub>A</sub>** und der **NASA-TLX Subskala geistige Anforderung** ( $r^{30} = .772, p < .001, 1 - \beta > .999$ ; Details s. Anhang H). Die RS(M)E<sub>A</sub>-Skala ermöglicht im Vergleich zu den Skalen des NASA-TLX eine feinere Differenzierung zwischen Personen hinsichtlich deren subjektiven (mental) Beanspruchungen aufgrund der höheren Skalenauflösung.

Die RSE<sub>A</sub> wird zur Erfassung der empfundenen allgemeinen Beanspruchung der MWB durch eine Tracking-Aufgabe verwendet. Die MWB erleben nicht nur eine subjektiv mentale Beanspruchung durch das Tracking, sondern auch körperliche Anstrengung durch die Joystick-Auslenkungen sowie sozialen Stress durch die kooperative Aufgabenbewältigung (Nachtwei, 2011b) und die Untersuchungssituation.

---

<sup>29</sup> Für eine systematische Ermittlung der Gütekriterien der RSME<sub>A</sub> und der RSE<sub>A</sub> sind bereits weitere Einsätze zusammen mit dem NASA-TLX (Grubert, n. D.) oder der SEA (Eilers et al., 1986) in Untersuchungen aus dem Bereich der Human Factors Forschung erfolgt respektive geplant. Die Ergebnisse werden gesondert veröffentlicht.

<sup>30</sup> Die Produkt-Moment-Korrelation nach Pearson ist das zur Ermittlung eines linearen Zusammenhangs zwischen intervallskalierten Variablen übliche Korrelationsmaß (Zöfel, 2003). Für den Korrelationskoeffizienten nach Pearson als deskriptives Maß liegen keine Verteilungsvoraussetzungen vor (Bortz, 1999, 2005). Allerdings setzt die inferenzstatistische Absicherung eine bivariate Normalverteilung der Variablen in der Population voraus. Der Erfüllung dieser Voraussetzung wird sich in der Praxis durch die Normalverteilung der zu korrelierenden Variablen angenähert. Diese ist jedoch weder für die RSME<sub>A</sub> noch die NASA-TLX-Skala „geistige Anforderung“ (Grubert, n. D.) gegeben. Zwar ist der Signifikanztest für die Produkt-Moment-Korrelation äußerst robust gegenüber Verletzungen der Verteilungsannahme (Bortz, 1999, 2005), dennoch wurde deshalb zusätzlich der Rangkorrelationskoeffizient tau von Kendall bestimmt:  $\tau = .560, p < .001, 1 - \beta > .999$ . Dieses Verfahren stellt keine Verteilungsvoraussetzungen an die Daten der zu korrelierenden Variablen. Die Rangkorrelation tau von Kendall kam in der gesamten Dissertation immer dann zum Einsatz, wenn bei mindestens einer der zu korrelierenden Variablen keine Normalverteilung vorlag. Somit wurde dieses Verfahren der, in der Literatur häufig zu findenden, Rangkorrelation nach Spearman vorgezogen. Aus Sicht der Autorin stellt es das geeignetere Verfahren dar, auch wenn mit kleineren Korrelationskoeffizienten zu rechnen ist. So geht der Rangkorrelationskoeffizient tau von Kendall nicht von äquidistanten Merkmalsabständen zwischen den aufeinanderfolgenden Rangplätzen aus (Bortz & Lienert, 1998). Zudem ist dieses Verfahren nicht ausreißersensibel (Bortz & Lienert, 1998). Ebenso stellt es das geeignetere Verfahren für kleine Stichproben und Rangbindungen dar (Field, 2005).

### 3.2.3.2.2 Fragebogen FESE-R

Für die Erhebung der unspezifischen Strategie wurde wie in den Voruntersuchungen der FESE-R von Nachtwei (2008b) verwendet (Details s. Anhang A.2.1).

### 3.2.3.2.3 Aufbau der Hauptuntersuchung

Der Untersuchungsaufbau von SAM 2.0 unterschied sich von dem von SAM 1.0 lediglich in den vorgenommenen Erweiterungen zur Komplexitätserhöhung der Versuchsumgebung (s. Abschnitt 3.1), wie der permanent sichtbaren leistungsabhängigen Entlohnungsstufen. Somit entspricht dieser dem der vier Voruntersuchungen (s. Abschnitt 3.2.1).

### 3.2.3.2.4 Experimentelles Design

Zur Aufklärung der Forschungsfragestellung 3, ob SAM 2.0 eine wahrnehmbare höhere Komplexität aufweist als SAM 1.0, wurden der Between-Faktor „SAM-Version“ (2fach gestuft: SAM 1.0 vs. 2.0) und der Within-Faktor „Durchgang“ (6fach gestuft: D7, D8, D10, D12, D14, D16) verwendet.

Die Untersuchung SAM 2.0 diente überdies der Überprüfung der Forschungsfragestellung 4, des Einflusses der Pfadlänge auf die wahrnehmbare Komplexität von SAM (Within-Faktor „Tracking-Pfadlänge“ 2fach gestuft: kTP und lTP). Dafür wurde zusätzlich ein 17. Tracking-Durchgang in SAM 2.0 integriert. Dieser ist kooperativ zu bewältigen und besitzt die dreifache Länge der anderen kooperativen Tracking-Durchgänge. Aus versuchsökonomischen Gründen wurde auf die Integration von weiteren kurzen und langen kooperativen Tracking-Durchgängen zur Berücksichtigung von Reihenfolge- und Positionseffekten verzichtet. Bei den subjektiven Beurteilungen der beiden Pfadversionen wurde zudem zwischen den beiden MWBN verglichen (Messwiederholungsfaktor „MWB“ 2fach gestuft: MWB 1 vs. MWB 2).

In SAM 2.0 wurden in die messrelevanten kooperativen Tracking-Durchgänge 7 bis 17 die beiden Gabelungen „elr“ und „rlr“, die beiden statischen Hindernisse „25lr“ und „50lr“ sowie das dynamische Hindernis implementiert. In die Tracking-Pfade der Durchgänge 7 bis 16 konnten jeweils drei Gabelungen, zwei statische Hindernisse und ein dynamisches Hindernis eingefügt werden (s. Tabelle 11). In den Tracking-Pfad des 17. Durchgangs konnten dreimal so viele Gabelungen und Hindernisse integriert werden. Die Gabelungen und Hindernisse wurden den einzelnen geeigneten Pfadstellen per Zufall zugeordnet, um die Varianz im Aufbau der Tracking-Pfade zu erhöhen (Details s. Anhang B.1.2).

Für die Aufklärung der Kontrollfragestellungen 1 bis 2 wurden zudem Kontrollvariablen von beiden MWBN als Selbsturteile (erneut der Messwiederholungsfaktor „MWB“ 2fach gestuft: MWB 1 vs. MWB 2) und teilweise als Beurteilungen des anderen MWBN (Within-

Faktor „MWB selbst vs. anderer“ 2fach gestuft: selbst vs. anderer) zu verschiedenen Zeitpunkten (Within-Faktor „Messzeitpunkt“ 6fach gestuft: MPrä1, MPrä2, MD8, MD12, MD15, MPost) erhoben (s. Tabelle 10).

Tabelle 11: Reihenfolge der Gabelungen und Hindernisse in den kooperativen Tracking-Durchgängen

Position im Tracking-Pfad	kooperativer Tracking-Durchgang						
	7	8	10	12	14	16	17
1	25lr	d	50lr	d	50lr	d	d
2	25lr	25lr	d	50lr	d	50lr	25lr
3	elr	elr	rll	elr	rll	elr	rll
4	rll	rll	elr	elr	rll	rll	elr
5	d	50lr	25lr	25lr	50lr	25lr	25lr
6	rll	rll	elr	elr	rll	elr	elr
7	Bei den Durchgängen 7 - 16 endete der Pfad nach der Position 6.  Die Positionen 7 bis 18 waren nur bei dem langen Pfad im letzten Durchgang gegeben.						50lr
8							d
9							elr
10							rll
11							50lr
12							rll
13							50lr
14							25lr
15							rll
16							elr
17							d
18							elr

Somit wurde ein sechsfaktorieller Versuchsplan (Mixed-Groups-Design, within & between) verwendet, wobei maximal zwei Faktoren zusammen in eine inferenzstatistische Analyse eingingen.

### 3.2.3.2.5 Durchführung der Hauptuntersuchung

Nachfolgend wird die Durchführung der Untersuchung von SAM 2.0 dargestellt.

#### 3.2.3.2.5.1 Probedurchläufe

Der Untersuchung von SAM 2.0 gingen zwei Probedurchläufe mit je einem VPn-Paar voraus, welche an den nachfolgenden Untersuchungen nicht mehr teilnahmen. Durch diese Durchläufe konnte die Angemessenheit der Untersuchungsplanung geprüft sowie der Ablauf eingeübt werden. Die Planung erwies sich als geeignet. Es mussten keine Änderungen und somit weitere Probedurchläufe durchgeführt werden.

### **3.2.3.2.5.2 Ablauf der Untersuchung**

Der Ablauf der Untersuchung von SAM 2.0 unterschied sich etwas von dem der vier Voruntersuchungen. Zum einen wurde den MWBn ein generelles Redeverbot ausgesprochen. Zum anderen bekamen diese nun zu Beginn auf dem Monitor fünf verschiedene Buttons zur gemeinschaftlichen Auswahl der vermeintlichen „Teamaufgabe“ präsentiert (s. Anhang B.1.1). Für den weiteren Verlauf der Untersuchung war es unerheblich für welchen der fünf Buttons sich ein Team entschied. Es folgten immer dieselben Instruktionen und derselbe Untersuchungsablauf. Des Weiteren wurden aufgrund der Erfahrungen aus den Voruntersuchungen die einzelnen Items teilweise in eine neue Reihenfolge mit einer anderen Einsatzhäufigkeit gestellt (s. Anhang A.2.4). Ferner mussten die MWB einen 17. Tracking-Durchgang absolvieren.

### **3.2.3.2.6 Untersuchungsteilnehmer**

Es folgt die Beschreibung der Akquise und Eigenschaften der VPn.

#### **3.2.3.2.6.1 Akquise**

Die Teilnehmerakquise für die Testung von SAM 1.0 und 2.0 erfolgte abermals über den Versuchspersonenserver PESA der Humboldt-Universität zu Berlin. Für beide Untersuchungen wurden die gleichen Ausschlusskriterien für die VPn wie in den Voruntersuchungen verwendet (s. Abschnitt 3.2.1.2.6). Einzig eine Vorselektion nach der unspezifischen Strategie fand in der Untersuchung SAM 2.0 nicht mehr statt.

#### **3.2.3.2.6.2 Beschreibung der Stichproben**

In der Untersuchung von SAM 1.0 waren 33 MWB-Paare eingesetzt worden. Davon ging ein MWB-Team aufgrund eines fehlenden Logfiles nicht mit in den Vergleich der beiden SAM-Versionen ein. Eine weitere Beschreibung dieser Stichprobe kann nicht vorgenommen werden, da Gross der Autorin aufgrund von Datenverlust die demografischen sowie Fragebogendaten der VPn nicht zur Verfügung stellen konnte.

Im Laufe der Untersuchung von SAM 2.0 kamen 39 MWB-Teams zum Einsatz. Bei einem MWB wurde der Fragebogen und bei zwei MWB-Paaren einzelne Tracking-Logfiles nicht automatisch gespeichert. Daher konnten nur die Daten von 36 MWB-Teams für den Vergleich mit der Untersuchung von SAM 1.0 herangezogen werden. Diese Stichprobe wurde bzgl. des Geschlechts als potentielle Störvariable ausbalanciert (18 Frauen- und 18 Männer-Teams). Das mittlere Alter lag bei 25.28 Jahren ( $SD = 4.21$ ). Die Stichprobe bestand hauptsächlich aus Studenten ( $H = 35$ ) gefolgt von Berufstätigen ( $H = 16$ ). Die Berufs- und Studienrichtungen

weisen eine hohe Spannbreite auf. Die einzelnen Richtungen sind, bis auf das Studium der Mathematik ( $H = 5$ ) und Informatik ( $H = 3$ ), nur ein- bis zweimal vertreten. (Details zur Stichprobe SAM 2.0 s. Anhang B.4.1)

### **3.2.3.3 Ergebnisse**

In diesem Kapitel werden die Aufbereitungen und die sich anschließenden Analysen der, in den Untersuchungen von SAM 1.0 und 2.0 erhobenen, objektiven Tracking-Daten und subjektiven Fragebogendaten vorgestellt. Die Ergebnisse werden in der Reihenfolge der Fragestellungen und Hypothesen präsentiert. Die Fragebogendaten aus der Untersuchung von SAM 1.0 konnten der Autorin nicht zur Verfügung gestellt werden. Daher wird nachfolgend nicht auf diese eingegangen. Ebenso konnte dadurch der Einfluss der Kontrollvariablen nicht aus den Varianzen der  $aV_n$  herauspartialisiert werden.

#### **3.2.3.3.1 Aufbereitung der Daten**

Nachfolgend werden zuerst die Aufbereitungen der Tracking-Daten und dann der Fragebogendaten beschrieben.

##### **3.2.3.3.1.1 Objektive Tracking-Daten**

Der Komplexitätsvergleich von SAM 1.0 und 2.0 erfolgte über die Leistung der MWB-Teams bei der Objektregelung. Diese Leistung wurde über die Daten der  $aV_n$  „benötigte Zeit“ und „Flächenfehler“ beim Tracking der kooperativen Tracking-Durchgänge 7 bis 16 mit Gabelungen und Hindernissen operationalisiert. Zudem wurden die Daten dieser  $aV_n$  vom Tracking-Durchgang 17 von SAM 2.0 für den Vergleich ein- vs. dreifache Pfadlänge verwendet.

Pro Tracking-Durchgang wurden die relevanten Untersuchungsdaten ca. alle 39 ms gemessen und als eine Datenzeile in ein XML-Logfile geschrieben. Für die Berechnung des Flächenfehlers aus diesem Logfile existierte bis dato nur ein zeit-, ressourcenintensives sowie fehleranfälliges Vorgehen. Deshalb war bei den vier Voruntersuchungen SAM 1.5 als Fehlermaß der zeitlich weniger aufwendig ermittelbare NrSensor verwendet worden. Für die Hauptuntersuchung entwickelte Hampel (2012) das effizientere Analyse-Tool „LogfileAnalyse 1.0“ (LFA 1.0). Dieses wurde hinsichtlich der Funktionalität und Gebrauchstauglichkeit von der Autorin mit konzipiert und auf Fehler getestet.

Das Tool ermittelt pro MWB-Paar und Tracking-Durchgang u. a. den Flächenfehler und die benötigte Zeit. Die Ergebnisse der Datenaufbereitung werden in einer XLS-Datei abgespeichert. Das LFA 1.0 wurde in Java programmiert und flexibel gestaltet, um notwendige Anpassungen im weiteren Verlauf des Projekts oder zur Nachnutzbarkeit von SAM für andere

Forscherguppen aufwandsgering vornehmen zu können. (s. Hampel, 2012)

Wie sich herausstellte, weisen die beiden SAM-Versionen unterschiedliche Softwareleistungen auf. Diese spiegeln sich in unterschiedlich lang andauernden Programmschritten wider.<sup>31</sup> Die mittlere Dauer eines Programmschritts in SAM 1.0 liegt bei 31.51 ms. In SAM 2.0 wird ein Programmschritt durchschnittlich nur alle 38.62 ms umgesetzt. Diese geringfügig schlechtere Softwareleistung von SAM 2.0 wird durch die vorgenommenen Erweiterungen verursacht. Trotz der Verwendung rechenstärkerer Computer und umfangreichen Maßnahmen zur Steigerung der Softwareleistung (s. Dullat, n. D.) konnte SAM 2.0 die Leistung von SAM 1.0 nicht erreichen.<sup>32</sup> Ein Programmschritt von SAM 1.0 benötigt somit im Mittel 18.41 % weniger Zeit als von SAM 2.0. Dadurch können die MWB das Objekt unter SAM 1.0 schneller regeln als unter SAM 2.0. Zudem weisen die Tracking-Durchgänge von SAM 1.0 und 2.0, aufgrund der unterschiedlichen Längen der verwendeten Gabelungen, unterschiedlich lange Tracking-Pfade auf. In SAM 1.0 ist ein Pfad 42483 Pixel lang, wohingegen ein Pfad in SAM 2.0 nur 39224 Pixeln entspricht. Damit ist ein Pfad in SAM 1.0 um 7.61 % länger als in SAM 2.0. Um trotz der unterschiedlichen Softwareleistungen und Tracking-Pfadlängen der Versionen „SAM 1.0 und 2.0“ hinsichtlich der aVn „benötigte Zeit“ und „Flächenfehler“ pro Tracking-Durchgang einen fairen Vergleich vornehmen zu können, wurden die Daten von SAM 1.0 an die Daten von SAM 2.0 angepasst. Dafür wurde zum einen von den Flächenfehlerdaten aus der Untersuchung von SAM 1.0 die 7.61 % subtrahiert. Dagegen wurde auf die Daten der benötigten Zeit in der Untersuchung SAM 1.0 die 10.80 % (18.41 % minus 7.61 %) addiert. Durch diese Transformierungsprozeduren beziehen sich diese angepassten Daten von SAM 1.0 ebenfalls auf eine durchschnittliche Dauer eines Programmschritts von 38.62 ms sowie eine Pfadlänge von 39224 Pixeln.

Für einen sinnvollen Vergleich des kurzen mit dem langen kooperativen Tracking-Pfad in SAM 2.0 wurden die objektiven Tracking-Daten der aVn „benötigte Zeit“ und „Flächenfehler“ beim Tracking des Durchgangs 17 an die Länge des kurzen Pfads angepasst. Der lange Pfad ist mit 116964 Pixeln 66.47 % länger als der kurze Pfad mit seinen 3922 Pixeln. Deshalb wurden von den Daten vom langen Pfad die 66.47 % subtrahiert. Dadurch beziehen sich die zu vergleichenden aVn auf die gleiche Pfadlänge.

---

<sup>31</sup> Zur Ermittlung der durchschnittlichen Dauer eines Programmschritts in SAM 1.0 und 2.0 diente die durchschnittlich benötigte Zeit für einen Programmschritt von jeweils per Zufall gewählten fünf unterschiedlichen Tracking-Durchgängen von jeweils einem anderen MWB-Team (s. Anhang B.2).

<sup>32</sup> Das Ziel für den Einsatz von Operateuren und Automatiken ist es, einen Programmschritt alle 30 ms umzusetzen.

Um die nachfolgende Ergebnispräsentation übersichtlich zu gestalten, werden als Einheiten für die benötigte Tracking-Zeit Sekunden und für den Flächenfehler Millionen Pixel verwendet.

### **3.2.3.3.1.2 Subjektive Fragebogendaten**

Die Erhebung der subjektiven Daten in der Untersuchung SAM 2.0 erfolgte mittels des FB-Tools des Lehrstuhls Ingenieurpsychologie/ Kognitive Ergonomie der Humboldt-Universität zu Berlin. Dieses Tool ermöglicht die Generierung eines Outputs der erhobenen Daten je Untersuchung in einer XLS-Datei. In diesem entspricht jede Zeile einer VPn und jede Spalte einem Item der Befragung. Aufgrund des Tool-Einsatzes auf zwei Laptops für MWB 1 und 2 entstanden zwei XLS-Dateien. Diese wurden manuell von der Autorin in eine XLS-Datei zusammengefügt. Die Daten der Abschlussbefragung per Papier wurden per Hand in diese XLS-Datei integriert.

### **3.2.3.3.2 Analyse der Daten**

Die Auswertung der aufbereiteten Rohdaten erfolgte mittels MS Office Excel 2007, SPSS 19 und GPower 3.1. Die deskriptiven Kennwerte (M, SD, H, %) zur Aufklärung der Forschungsfragestellung 3 und 4 sind dem Anhang B.4.2 und der Kontrollfragestellung 1 und 2 sind dem Anhang B.4.3 zu entnehmen.

Für die inferenzstatistischen Absicherungen der Ergebnisse zur Überprüfung von Unterschieds- und Veränderungshypothesen wurde das Alpha-Fehler-Niveau a priori auf 5 %<sup>33</sup> festgelegt. Für die Überprüfung von Gleichheitshypothesen wurde das Alpha-Fehler-Niveau a priori auf 20 %<sup>34</sup> angehoben.

Für die Überprüfung der einzelnen Hypothesen wurden verschiedene statistische Verfahren verwendet. Wurde eine univariate Varianzanalyse mit Messwiederholung<sup>35, 36</sup> bei nicht

---

<sup>33</sup> Dadurch wird die Wahrscheinlichkeit einer Fehlentscheidung zugunsten der Alternativhypothese (Unterschieds-/ Veränderungshypothese) klein gehalten (Bortz, 1999).

<sup>34</sup> Dadurch wird die Wahrscheinlichkeit einer Fehlentscheidung zugunsten der Nullhypothese (Gleichheitshypothese) minimiert (Bortz, 1999).

<sup>35</sup> Nach Bortz (1999) ist die Varianzanalyse bei gleichgroßen Stichproben mit mehr als zehn VPn relativ robust gegenüber der Verletzung von deren Voraussetzung. Deshalb wurde trotz teilweise nicht normalverteilter Variablen die einfaktorielle Varianzanalyse der non-parametrischen Rangvarianzanalyse von Friedman vorgezogen. Ein nicht signifikantes Ergebnis unter Verletzung der Voraussetzung würde sich auch bei gegebener Normalverteilung zeigen (Zöfel, 2003).

<sup>36</sup> Es existiert kein äquivalentes non-parametrisches Verfahren zur mehrfaktoriellen Varianzanalyse. Nach Bortz (1999) und Zöfel (2003) handelt es sich jedoch bei der Varianzanalyse um ein ziemlich robustes Verfahren gegenüber den Verletzungen von deren Voraussetzungen. Ein nicht signifikantes Ergebnis unter Verletzung der Voraussetzungen würde sich auch bei gegebener Normalverteilung und Varianzhomogenität zeigen (Zöfel, 2003).



gegebener Normalverteilung<sup>37</sup> der Messwerte der einzelnen Zellen ausgeführt, wurde mit einem Alpha-Fehler-Niveau von 4 %<sup>38</sup> getestet. Weisen die Stichproben keine Varianzhomogenität<sup>39</sup> auf, wurde die Signifikanzschranke auf 1 %<sup>40</sup> gesetzt.

Für Messwiederholungsfaktoren mit mehr als zwei Stufen wurde die Sphärizitätsvoraussetzung mittels des Mauchly-Tests geprüft. Bei Vorliegen inhomogener Varianzen der Differenzen der Messungen zwischen dessen Stufen wurden korrigierte Freiheitsgrade für den kritischen F-Wert genutzt. Dabei wurde der Empfehlung von Field (2005) gefolgt bei einem  $\epsilon < .75$  das Greenhouse-Geisser Verfahren und bei einem  $\epsilon > .75$  das Huynh-Feldt Verfahren zu verwenden. Aus Platzgründen werden nur die hypothesenrelevanten Ergebnisse der Varianzanalysen präsentiert.

Für einen Vergleich der Mittelwerte zweier unabhängiger Stichproben wurde der t-Test für unabhängige Stichproben eingesetzt. Voraussetzung dafür ist eine Normalverteilung der Messwerte beider Stichproben bzgl. der relevanten Variable. Liegt bei mindestens einer der beiden Stichproben keine Normalverteilung vor, wurde das non-parametrische Pendant, der Mann-Whitney-U-Test, eingesetzt.

Für einen Vergleich der Mittelwerte zweier abhängiger Stichproben wurde der t-Test für abhängige Stichproben eingesetzt. Voraussetzung dafür ist bei Stichproben von mehr als 30 VPn eine signifikante positive Korrelation<sup>41</sup> zwischen diesen bzgl. der relevanten Variable.<sup>42</sup> Bei zwei Stichproben mit weniger als 30 VPn muss zudem eine Normalverteilung der Differenzen dieser beiden Stichproben hinsichtlich der relevanten Variable gegeben sein. Sind diese Bedingungen nicht erfüllt, wurde der Vorzeichenrangtest von Wilcoxon genutzt.

---

<sup>37</sup> Für die Überprüfung der Normalverteilung kam bei einer Stichprobe von mehr als 50 VPn nur der Kolmogorov-Smirnov-Test mit Lillifors-Schranken zum Einsatz. Bei kleineren Stichproben wurden sowohl der Kolmogorov-Smirnov-Test mit Lillifors-Schranken als auch der Shapiro-Wilk-Test verwendet. Der Shapiro-Wilk-Test ist für Stichproben unter 50 Personen ausgelegt. Für  $N < 50$  weist dieser eine hohe Teststärke auf (Seier, 2002). Die Normalverteilung einer Variable wurde in diesen Fällen nur angenommen, wenn diese durch beide Verfahren mit einem nicht signifikanten Ergebnis ( $p > .20$ ) statistisch bestätigt werden konnte.

<sup>38</sup> Nach Zöfel (2003) sollte bei nicht gegebener Normalverteilung die Signifikanzgrenze von .05 auf .04 gesetzt werden.

<sup>39</sup> Die Varianzhomogenität wurde mittels des Levene-Tests überprüft.

<sup>40</sup> Bühl und Zöfel (2005) empfehlen bei nicht gegebener Varianzhomogenität die Signifikanzgrenze von .05 auf .01 zu setzen.

<sup>41</sup> Bei gegebener Normalverteilung der relevanten Variable bei beiden Stichproben wurde der Korrelationskoeffizient nach Pearson ermittelt. Lag bei mindestens einer der beiden Stichproben bzgl. der relevanten Variable keine Normalverteilung vor, wurde der Rangkorrelationskoeffizient nach Kendall berechnet.

<sup>42</sup> Laut Bortz (1999) ist der t-Test für abhängige Stichproben bei mehr als 30 VPn relativ robust gegenüber der Voraussetzungsverletzung der Normalverteilung der Differenzen der beiden Stichproben hinsichtlich der relevanten Variable. Es sollte dann jedoch eine positive Korrelation zwischen den beiden Stichproben bzgl. der relevanten Variable vorliegen, um einen Verlust an Teststärke vorzubeugen.

Liegen Ausreißer<sup>43</sup> vor, wurden die deskriptiven und inferenzstatistischen Analysen mit und ohne Ausreißer durchgeführt. Nachfolgend werden die Ergebnisse inklusive Ausreißer berichtet, da Messfehler in den Daten aufgrund der Erfassungsmethoden ausgeschlossen werden.<sup>44</sup> Ergab eine statistische Überprüfung unter Ausschluss der Ausreißer jedoch ein bedeutsam anderes Ergebnis, wird dieses ebenfalls berichtet und diskutiert.

### 3.2.3.3.3 Ergebnisse der Analysen

Nachfolgend werden die Ergebnisse zur Aufklärung der Forschungsfragestellung 3 und 4 sowie der Kontrollfragestellungen 1 und 2 in dieser Reihenfolge präsentiert.

#### 3.2.3.3.3.1 Ergebnisse der Analysen zur Forschungsfragestellung 3

Als erstes wurde überprüft, ob zwischen den objektiven Tracking-Daten der SAM-Versionen 1.0 und 2.0, als Indikatoren der wahrnehmbaren Komplexität, Unterschiede vorliegen. Zudem wurde explorativ die zeitliche Stabilität dieser aVn betrachtet. Daher wurden für die statistischen Vergleiche zweifaktorielle univariate Varianzanalysen mit dem Gruppierungsfaktor „SAM-Version“ und dem Messwiederholungsfaktor „Durchgang“ ausgeführt.<sup>45</sup> Bei einem globalen signifikanten Unterschied zwischen den beiden SAM-Versionen wurde für jeden der sechs Durchgänge ein Einzelvergleich mittels des t-Tests für unabhängige Stichproben bzw. des Mann-Whitney-U-Tests zwischen SAM 1.0 und 2.0 hinsichtlich der entsprechenden aV durchgeführt. Aufgrund der Alpha-Fehler-Kumulierung<sup>46</sup> durch die sechs Einzelvergleiche je aV wurde eine Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .0083$  vorgenommen.

### 3.1 Forschungshypothese bezüglich aV „Genauigkeit beim Objekt-Tracking“

Die *Hypothese F.3.1* postuliert einen gerichteten Unterschied hinsichtlich der Anzahl der Fehler beim Tracking zwischen den beiden SAM-Versionen. Unter Einsatz von SAM 2.0 produzieren die MWB mehr Fehler als bei SAM 1.0. Tabelle 12 präsentiert den durchschnittlich

---

<sup>43</sup> Ausreißeranalysen erfolgten bei normalverteilten Variablen mittels einer z-Wert-Analyse und bei nicht normalverteilten Variablen über den Boxplot.

<sup>44</sup> Um im eingesetzten FB-Tool auf den Laptops der beiden MWB von einer zur nächsten Frage blättern zu können, mussten die MWB eine Beantwortung vornehmen. So existierte bei den Radio-Buttons und Checkboxes keine Vorauswahl, die Schieberegler befanden sich auf der Nullstellung. Damit sind ungewollte Antworten weitestgehend auszuschließen.

<sup>45</sup> Durch die Aufnahme eines weiteren Faktors in eine Varianzanalyse kann die Teststärke für die statistische Überprüfung des Effekts eines Faktors erhöht werden. Dafür muss jedoch die aufgeklärte Variation der aV durch den zusätzlich aufgenommenen Faktor den jeweiligen Verlust an Freiheitsgraden übersteigen. Ferner ermöglicht jede weitere Stufe des zusätzlichen Faktors die Erhöhung der statistischen Power eines Tests für den überprüften Effekt eines Faktors. (Sedlmeier & Renkewitz, 2008)

<sup>46</sup> Laut Bortz (1999) verursacht die Überprüfung einer Hypothese durch mehrere simultane Signifikanztests eine Erhöhung der Wahrscheinlichkeit die Nullhypothese fälschlicherweise zu verwerfen. Dieser Fehler wird durch eine Alpha-Fehler-Adjustierung vermieden.

produzierten Flächenfehler der MWB je relevantem Durchgang der Untersuchungen von SAM 1.0 und 2.0.

Tabelle 12: Mittelwert und Standardabweichung des Flächenfehlers beim Tracking in Millionen Pixel pro Durchgang und über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0

	SAM 1.0		SAM 2.0		Gesamt (SAM 1.0 & 2.0)	
	M	SD	M	SD	M	SD
Fehler Durchgang 7 in Mio px	1.03	0.25	1.78	0.26	1.43	0.45
Fehler Durchgang 8 in Mio px	0.96	0.21	1.83	0.23	1.42	0.49
Fehler Durchgang 10 in Mio px	1.01	0.23	1.75	0.21	1.40	0.43
Fehler Durchgang 12 in Mio px	0.97	0.21	1.72	0.30	1.36	0.46
Fehler Durchgang 14 in Mio px	1.01	0.22	1.79	0.32	1.42	0.48
Fehler Durchgang 16 in Mio px	0.95	0.27	1.75	0.25	1.37	0.48
Fehler gesamt in Mio px	0.99	0.23	1.77	0.26		

Anmerkungen. M = Mittelwert; SD = Standardabweichung; Mio px = Millionen Pixel

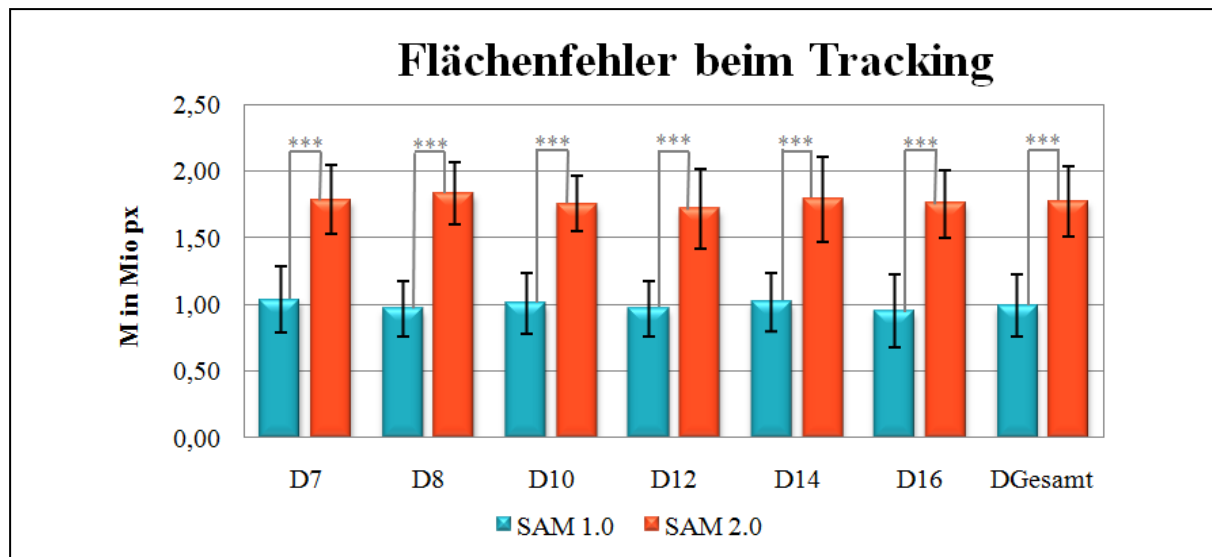


Abbildung 9: Mittelwert (M) und Standardabweichung des Flächenfehlers beim Tracking in Millionen Pixel (Mio px) pro Durchgang (D) sowie über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0

Anmerkungen: signifikante Unterschiede sind markiert: \*\*\* =  $p_{\text{einseitig}} < .001$  respektive die Alpha-Fehler-Korrektur dieser Signifikanzschranke

### 3.2 Forschungshypothese bezüglich $\alpha V$ „Geschwindigkeit beim Objekt-Tracking“

Die **Hypothese F.3.2** postuliert, dass unter Einsatz von SAM 2.0 weniger Zeit zum Tracking benötigt wird als bei SAM 1.0. Tabelle 13 präsentiert die durchschnittlich benötigte Tracking-Zeit je relevantem Durchgang der Untersuchungen von SAM 1.0 und 2.0.

Für den Gruppierungsfaktor „**SAM-Version**“ (2fach gestuft) zeigte sich unter Einschluss der Ausreißer bei einseitiger Testung eine Overall-Signifikanz bzgl. der benötigten Zeit zum Tracking ( $F[1, 66] = 5.863$ ,  $p_{\text{zweiseitig}} = .018$ ,  $\eta^2_p = .081$ ,  $1 - \beta = .88$ ,  $p_{\text{einseitig}} = .009$ ; s. Abbildung 10). Bei den zusätzlich ausgeführten sechs Einzelvergleichen liegen bei einseitigen

Testungen diesbzgl. lediglich bei drei Durchgängen Tendenzen<sup>47</sup> zu einem signifikanten Unterschied zwischen den beiden SAM-Versionen vor (**D7**:  $U = 393.00$ ,  $p_{\text{einseitig}} = .012$ ,  $r = -.273$ ,  $1 - \beta = .42$ ; **D8**:  $U = 394.00$ ,  $p_{\text{einseitig}} = .013$ ,  $r = -.271$ ,  $1 - \beta = .44$ ; **D12**:  $U = 403.00$ ,  $p_{\text{einseitig}} = .017$ ,  $r = -.258$ ,  $1 - \beta = .47$ ).

Tabelle 13: Mittelwert und Standardabweichung der benötigten Zeit zum Tracking in Sekunden pro Durchgang und über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0

	SAM 1.0		SAM 2.0		Gesamt (SAM 1.0 & 2.0)	
	M	SD	M	SD	M	SD
<b>mit Ausreißer</b>						
Zeit Durchgang 7 in s	225.92	76.45	188.13	53.60	205.91	67.58
Zeit Durchgang 8 in s	218.69	72.26	182.13	49.91	199.34	63.68
Zeit Durchgang 10 in s	211.12	72.43	175.90	44.32	192.47	61.38
Zeit Durchgang 12 in s	207.79	70.17	171.71	44.56	188.69	60.37
Zeit Durchgang 14 in s	197.56	64.72	167.47	42.28	181.63	55.68
Zeit Durchgang 16 in s	184.97	69.40	165.91	40.38	174.87	56.32
Zeit Gesamt in s	207.67	71.34	175.21	46.21		
<b>ohne Ausreißer</b>						
Zeit Durchgang 7 in s	218.94	66.54	173.43	34.41	195.82	57.10
Zeit Durchgang 8 in s	212.44	64.06	168.95	33.30	190.35	54.96
Zeit Durchgang 10 in s	205.89	67.22	164.77	29.47	185.00	55.23
Zeit Durchgang 12 in s	205.59	70.20	160.23	29.71	182.55	57.86
Zeit Durchgang 14 in s	194.43	63.29	157.43	28.89	175.63	51.99
Zeit Durchgang 16 in s	190.89	61.78	156.58	27.06	173.46	50.12
Zeit Gesamt in s	204.70	65.40	163.56	30.80		

Anmerkungen. M = Mittelwert; SD = Standardabweichung; s = Sekunde

Bei Ausschluss der fünf gefundenen Ausreißer ergeben sich jedoch bei den vier Durchgängen 7, 8, 10 und 12 signifikante Unterschiede zwischen **SAM 1.0** und **2.0** (**D7**:  $U = 286.00$ ,  $p_{\text{einseitig}} = .002$ ,  $r = -.364$ ,  $1 - \beta = .72$ ; **D8**:  $U = 290.00$ ,  $p_{\text{einseitig}} = .002$ ,  $r = -.357$ ,  $1 - \beta = .72$ ; **D10**:  $U = 319.00$ ,  $p_{\text{einseitig}} = .007$ ,  $r = -.307$ ,  $1 - \beta = .66$ ; **D12**:  $U = 305.00$ ,  $p_{\text{einseitig}} = .004$ ,  $r = -.331$ ,  $1 - \beta = .71$ ). Bei dem Durchgang 14 liegt diesmal zumindest eine Tendenz zur Signifikanz vor ( $U = 328.00$ ,  $p_{\text{einseitig}} = .010$ ,  $r = -.291$ ,  $1 - \beta = .61$ ). Lediglich beim Durchgang 16 zeigt sich erneut kein signifikantes Ergebnis ( $U = 349.00$ ,  $p_{\text{einseitig}} = .022$ ,  $r = -.255$ ,  $1 - \beta = .57$ ).

Für den Messwiederholungsfaktor „Durchgang“ (6fach gestuft) ergibt sich ein höchst signifikanter Einfluss auf die benötigte Zeit zum Tracking ( $F[1.630, 107.609]^{48} = 16.845$ ,  $p <$

<sup>47</sup> Aufgrund der Alpha-Fehler-Kumulierung durch die sechs Einzelvergleiche liegt die Schranke der Tendenz einer Signifikanz bei  $p = .017$ .

<sup>48</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Greenhouse-Geisser Verfahrens korrigiert.

.001,  $\eta^2_p = .203$ ,  $1 - \beta > .999$ ). Zur Ermittlung, ob dieses Ergebnis sowohl durch SAM 1.0 als auch 2.0 verursacht wurde, wurde je Version eine einfaktorielle univariate Varianzanalyse mit dem Messwiederholungsfaktor „Durchgang“ (6fach gestuft) durchgeführt. Der Alpha-Fehler-Kumulierung wurde hierbei mit einer Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .025$  entgegengewirkt.

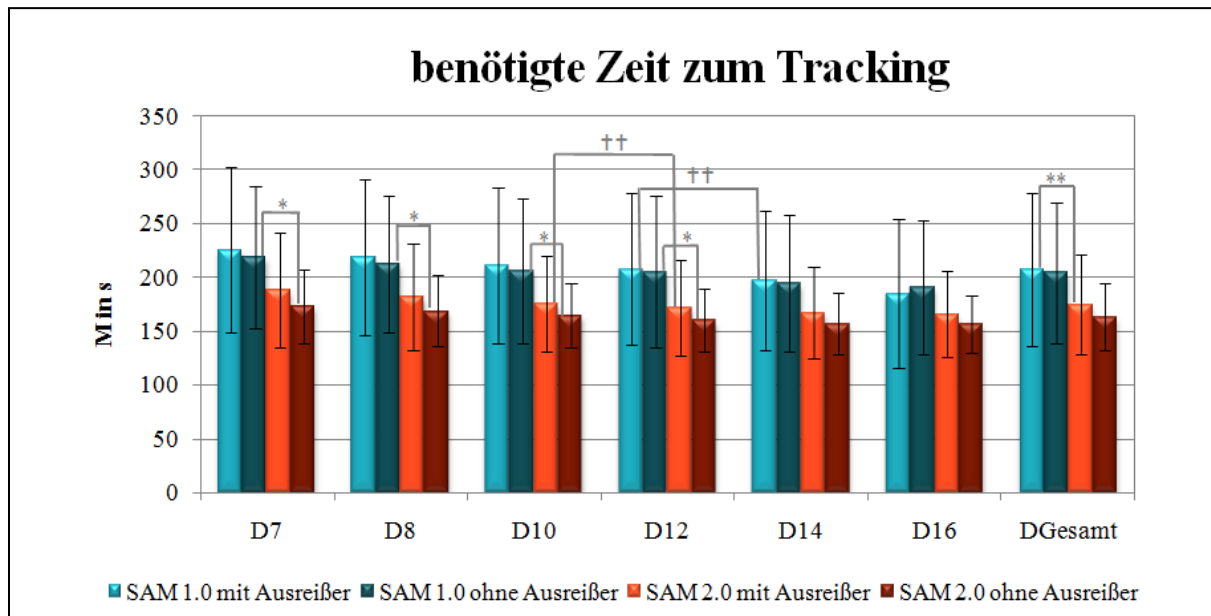


Abbildung 10: Mittelwert (M) und Standardabweichung der benötigten Zeit zum Tracking in Sekunden (s) mit und ohne Ausreißer pro Durchgang (D) sowie über alle Durchgänge der Untersuchungen von SAM 1.0 und 2.0

*Anmerkungen.* Signifikante Unterschiede sind markiert: \* =  $p_{\text{einseitig}} < .05$ , \*\* =  $p_{\text{einseitig}} < .01$  und †† =  $p_{\text{zweiseitig}} < .01$  respektive die jeweilige Alpha-Fehler-Korrektur dieser drei Signifikanzschranken. Für die Daten ohne Ausreißer werden, zum Erhalt der Übersicht, diese Markierungen nur vorgenommen, wenn sich die Signifikanzwerte von denen der Daten mit Ausreißer unterscheiden.

Bei **SAM 1.0** liegt für den Faktor „**Durchgang**“ ebenfalls ein sehr signifikanter Unterschied bzgl. der benötigten Zeit zum Tracking vor ( $F[1.473, 45.673]^{49} = 7.212$ ,  $p = .004$ ,  $\eta^2_p = .189$ ,  $1 - \beta > .999$ ). Zur Überprüfung, wann es im Verlauf der Untersuchung zu einer signifikanten Veränderung hinsichtlich der benötigten Zeit zum Tracking kam, wurden nur Einzelvergleiche zwischen den aufeinanderfolgenden Durchgängen durchgeführt. Die Unterschiede zwischen diesen Durchgängen spiegeln die zeitlichen Veränderungen bei dieser aV am deutlichsten wider. Auf die weiteren möglichen Einzelvergleiche wurde verzichtet, um die Alpha-Fehler-Kumulierung möglichst klein zu halten. Aufgrund der fünf ausgeführten Einzelvergleiche, wurde erneut eine Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .01$  angewendet. Bei den fünf Einzelvergleichen mittels des t-Tests für

<sup>49</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Greenhouse-Geisser Verfahrens korrigiert.

abhängige Stichproben ergab sich nur für „**D12 vs. D14**“ ein signifikantes Ergebnis ( $SE_{D12} = 12.40$ ,  $SE_{D14} = 11.44$ ,  $t[31] = 3.605$ ,  $p = .001$ ,  $r = .543$ ,  $1 - \beta = .80$ ). Bei dem Einzelvergleich „D8 vs. D10“ zeigte sich eine Tendenz<sup>50</sup> zur Signifikanz ( $SE_{D8} = 12.77$ ,  $SE_{D10} = 12.80$ ,  $t[31] = 2.613$ ,  $p = .014$ ,  $r = .425$ ,  $1 - \beta = .46$ ). Bei den anderen drei Einzelvergleichen „D7 vs. D8“, „D10 vs. D12“ und „D14 vs. D16“ fanden sich keine signifikanten Unterschiede („D7 vs. D8“:  $SE_{D7} = 13.52$ ,  $SE_{D8} = 12.77$ ,  $t[31] = 1.538$ ,  $p = .134$ ,  $r = .266$ ,  $1 - \beta = .13$ ; „D10 vs. D12“:  $SE_{D10} = 12.80$ ,  $SE_{D12} = 12.40$ ,  $t[31] = 0.867$ ,  $p = .392$ ,  $r = .154$ ,  $1 - \beta = .04$ ; „D14 vs. D16“:  $SE_{D14} = 11.44$ ,  $SE_{D16} = 12.27$ ,  $t[31] = 1.355$ ,  $p = .185$ ,  $r = .236$ ,  $1 - \beta = .10$ ).

Bei **SAM 2.0** ergibt sich ebenso für den Faktor „**Durchgang**“ ein höchst signifikanter Unterschied hinsichtlich der benötigten Zeit zum Tracking ( $F[1.615, 56.673]^{51} = 16.976$ ,  $p < .001$ ,  $\eta^2_p = .327$ ,  $1 - \beta > .999$ ). Bei den fünf Einzelvergleichen mittels des t-Tests für abhängige Stichproben fand sich nur für „**D10 vs. D12**“ ein signifikantes Ergebnis ( $SE_{D10} = 7.39$ ,  $SE_{D12} = 7.43$ ,  $t[35] = 3.571$ ,  $p = .001^{52}$ ,  $r = .517$ ,  $1 - \beta = .80$ ). Eine Tendenz<sup>53</sup> zur Signifikanz zeigten die drei Einzelvergleiche „D7 vs. D8“, „D8 vs. D10“ und „D12 vs. D14“ auf („D7 vs. D8“:  $SE_{D7} = 8.93$ ,  $SE_{D8} = 8.32$ ,  $t[35] = 2.491$ ,  $p = .018$ ,  $r = .388$ ,  $1 - \beta = .42$ ; „D8 vs. D10“:  $SE_{D8} = 8.32$ ,  $SE_{D10} = 7.39$ ,  $t[35] = 2.590$ ,  $p = .014$ ,  $r = .401$ ,  $1 - \beta = .46$ ; „D12 vs. D14“:  $SE_{D12} = 7.43$ ,  $SE_{D14} = 7.05$ ,  $t[35] = 2.580$ ,  $p = .014$ ,  $r = .400$ ,  $1 - \beta = .45$ ). Bei dem Einzelvergleich „D14 vs. D16“ ergab sich kein signifikanter Unterschied ( $SE_{D14} = 7.05$ ,  $SE_{D16} = 6.73$ ,  $t[35] = 1.172$ ,  $p = .249$ ,  $r = .194$ ,  $1 - \beta = .07$ ).

### 3.2.3.3.2 Ergebnisse der Analysen zur Forschungsfragestellung 4

Für die Überprüfung der Forschungsfragestellung 4, bei welcher Pfadlänge SAM eine wahrnehmbare höhere Komplexität aufweist, wurden die beiden unterschiedlich langen, kooperativ zu absolvierenden, Tracking-Pfade hinsichtlich der aVn „Flächenfehler“ und „benötigte Zeit“ beim Tracking sowie „subjektive Beanspruchung“ und „subjektive Komplexität“ der Tracking-Aufgabe verglichen. Damit Zeit- und Lerneffekte diese Vergleiche so wenig wie möglich beeinflussen, gingen in die statistischen Analysen nicht die Durchschnittswerte der Daten der aVn von allen kurzen kooperativ zu absolvierenden Tracking-Pfaden ein, sondern nur des Durchgangs 16.

<sup>50</sup> Aufgrund der Alpha-Fehler-Kumulierung durch die fünf Einzelvergleiche liegt die Schranke der Tendenz einer Signifikanz bei  $p = .02$ .

<sup>51</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Greenhouse-Geisser Verfahrens korrigiert.

<sup>52</sup> Aufgrund der fünf ausgeführten Einzelvergleiche wurde erneut eine Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .01$  angewendet.

<sup>53</sup> Aufgrund der Alpha-Fehler-Kumulierung durch die fünf Einzelvergleiche liegt die Schranke der Tendenz einer Signifikanz bei  $p = .02$ .

Für den Vergleich zwischen dem kurzen und dem langen Pfad hinsichtlich der verursachten Fehler beim Tracking (s. Tabelle 14) wurde ein t-Test für abhängige Stichproben verwendet. Es liegt kein signifikanter Unterschied zwischen den beiden Pfadlängen hinsichtlich dieser aV vor ( $t[35] = 0.501$ ,  $p = .619$ ,  $r = .084$ ,  $1 - \beta = .12$ ).

Tabelle 14: Mittelwert, Standardabweichung und Standardfehler des Mittelwerts des Flächenfehlers beim Tracking in Millionen Pixel für den kurzen und langen Tracking-Pfad

	Kurzer Tracking-Pfad			Langer Tracking-Pfad		
	M	SD	SE	M	SD	SE
Fehler in Mio px	1.75	0.25	0.042	1.74	0.23	0.039

Anmerkungen. M = Mittelwert; SD = Standardabweichung; SE = Standardfehler des Mittelwerts; Mio px = Millionen Pixel

Ein t-Test für abhängige Stichproben wurde ebenfalls für den Vergleich des **kurzen** und des **langen Pfads** bzgl. der **benötigten Zeit** zum Tracking eingesetzt. Es zeigt sich, dass beim langen Pfad signifikant weniger Zeit zum Tracking benötigt wird ( $t[35] = 2.125$ ,  $p = .041$ ,  $r = .338$ ,  $1 - \beta = .67$ ; s. Abbildung 11). Allerdings ist dieser Unterschied bei Ausschluss der zwei Ausreißer nicht mehr signifikant ( $t[33] = 1.586$ ,  $p = .122$ ,  $r = .266$ ,  $1 - \beta = .46$ ).

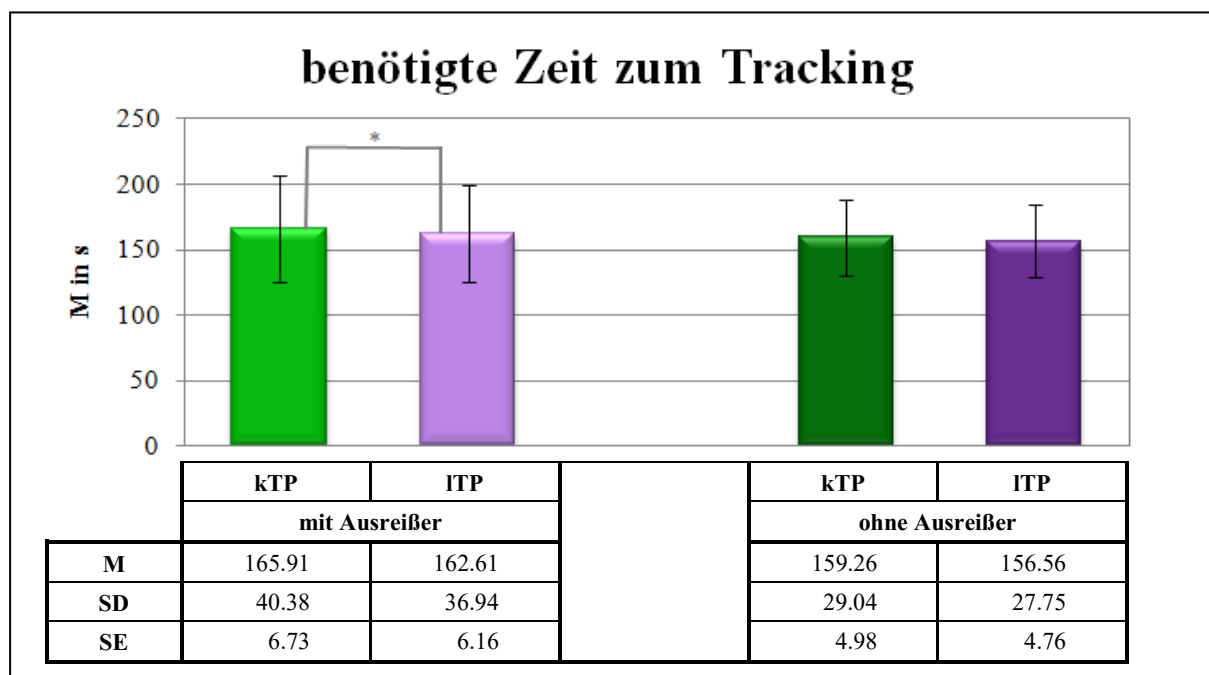


Abbildung 11: Mittelwert (M), Standardabweichung (SD) und Standardfehler des Mittelwerts (SE) der benötigten Zeit zum Tracking in Sekunden (s) mit und ohne Ausreißer für den kurzen (kTP) und langen (ITP) Tracking-Pfad

Anmerkung. signifikante Unterschiede sind markiert: \* =  $p_{\text{zweiseitig}} < .05$

Für den Vergleich der beiden Tracking-Pfade hinsichtlich der subjektiven Beanspruchung der MWB durch das Tracking wurde eine zweifaktorielle univariate Varianzanalyse mit Messwiederholung durchgeführt. Für den Messwiederholungsfaktor „Tracking-

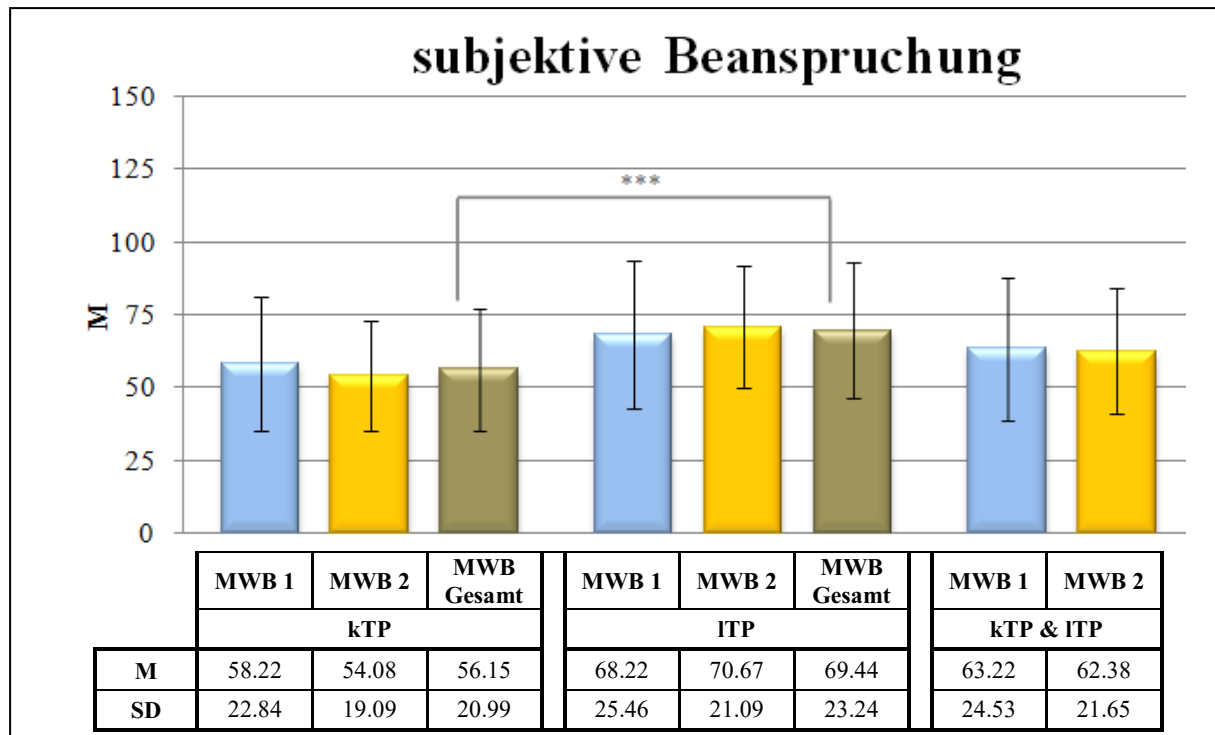


Abbildung 12: Mittelwert (M) und Standardabweichung (SD) der subjektiven Beanspruchung von MWB 1 und 2 nach der RSE<sub>A</sub> für den kurzen (kTP) und langen (ITP) Tracking-Pfad sowie über beide hinweg

*Anmerkungen.* Skala: 0 = absolut keine Anstrengung bis 150 = unerträgliche Anstrengung; signifikante Unterschiede sind markiert: \*\*\* =  $p_{\text{zweiseitig}} < .001$

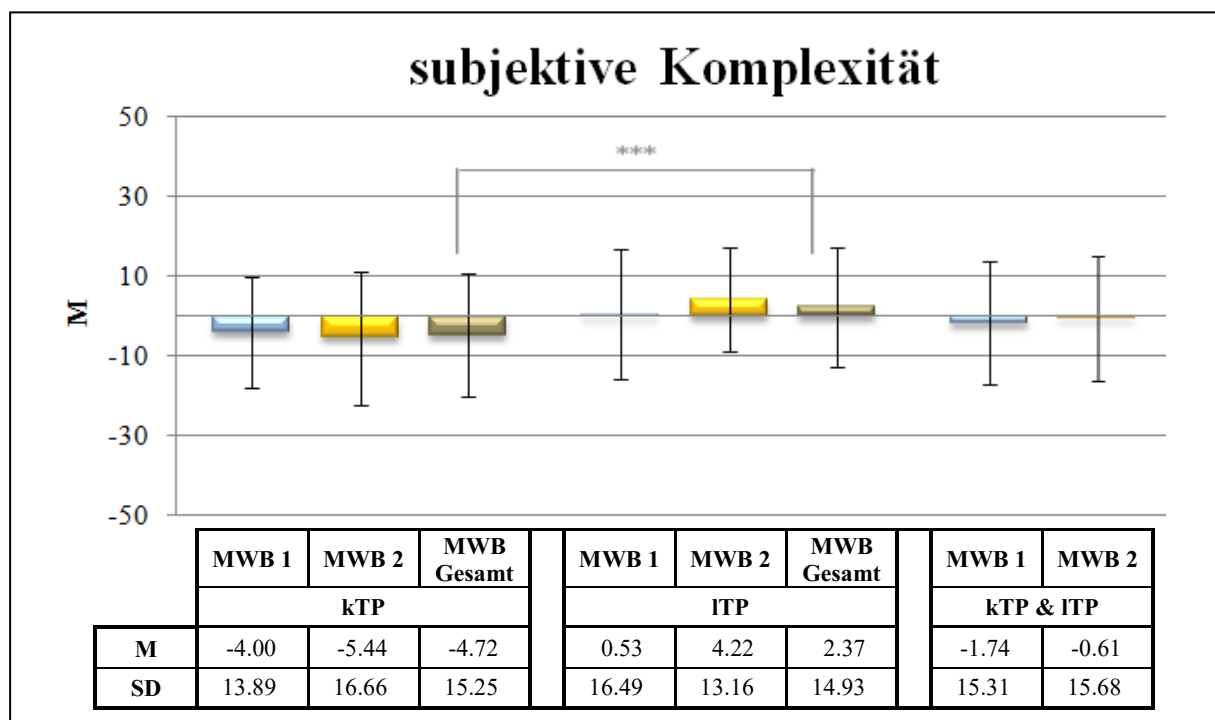


Abbildung 13: Mittelwert (M) und Standardabweichung (SD) der subjektiven Komplexität nach MWB 1 und 2 für den kurzen (kTP) und langen (ITP) Tracking-Pfad sowie über beide hinweg

*Anmerkungen.* Skala: -50 = niedrige Komplexität bis 50 = hohe Komplexität; signifikante Unterschiede sind markiert: \*\*\* =  $p_{\text{zweiseitig}} < .001$



**Pfadlänge**“ (2fach gestuft) ergibt sich ein höchst signifikanter Unterschied, wie anstrengend die MWB die Tracking-Aufgabe empfunden haben ( $F[1, 35] = 93.621$ ,  $p < .001$ ,  $\eta^2_p = .728$ ,  $1 - \beta > .999$ ; s. Abbildung 12). Für den Messwiederholungsfaktor „MWB“ (2fach gestuft) zeigt sich kein signifikanter Einfluss auf die subjektive Beanspruchung der MWB ( $F[1, 35] = 0.035$ ,  $p = .852$ ,  $\eta^2_p = .001$ ,  $1 - \beta = .07$ ).

Für den Vergleich der beiden Tracking-Pfade hinsichtlich der **subjektiven Komplexität** der Tracking-Aufgabe wurde erneut eine zweifaktorielle univariate Varianzanalyse mit Messwiederholung durchgeführt. Für den Messwiederholungsfaktor „**Tracking-Pfadlänge**“ (2fach gestuft) besteht eine höchst signifikant höhere Komplexität des langen Pfads ( $F[1, 35] = 25.510$ ,  $p < .001$ ,  $\eta^2_p = .422$ ,  $1 - \beta > .999$ ; s. Abbildung 13). Für den Messwiederholungsfaktor „MWB“ (2fach gestuft) zeigt sich kein signifikanter Einfluss auf die subjektive Komplexität der Tracking-Aufgabe ( $F[1, 35] = 0.111$ ,  $p = .741$ ,  $\eta^2_p = .003$ ,  $1 - \beta = .10$ ).

### 3.2.3.3.3 Ergebnisse der Analysen zur Kontrollfragestellung 1

Nachfolgend werden die Analyseergebnisse zur Aufklärung der Kontrollfragestellung 1 vorgestellt, ob die Rahmenbedingungen der Versuchsumgebung SAM 2.0 die MWB motivieren, deren spezifische Strategie beim Tracking umzusetzen.

#### 1.1 Kontrollhypothese bezüglich Instruktionsverständlichkeit

Die **Hypothese K.1.1** postuliert, dass die Instruktion der jeweiligen spezifischen Strategie für MWB 1 und 2 subjektiv verständlich ist und diesbzgl. kein Unterschied zwischen ihnen vorliegt.

Tabelle 15: Deskriptive Statistiken der Beurteilungen durch die MWB bzgl. Instruktionsverständlichkeit, Instruktionsbefolgung und Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen

	M	SD	Median	% der MWB mit SW > 10	% der MWB mit SW < -10
<b>Instruktionsverständlichkeit</b>					
<b>MWB 1</b>	39.78	14.95	48.00	91.70	0.00
<b>MWB 2</b>	42.00	14.21	50.00	94.40	0.00
<b>Instruktionsbefolgung</b>					
<b>MWB 1</b>	29.42	16.69	31.50	88.90	2.80
<b>MWB 2</b>	29.17	20.55	37.50	80.60	5.60
<b>Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen</b>					
<b>MWB 1</b>	26.03	27.31	29.50	80.60	8.30
<b>MWB 2</b>	28.58	18.74	30.00	80.60	0.00

Anmerkungen. Skala: -50 = gar nicht verständlich/ gar nicht befolgt/ unwichtig bis 50 = völlig verständlich/ vollends befolgt/ sehr wichtig; M = Mittelwert; SD = Standardabweichung; % = Prozent; SW = Skalenwert

Tabelle 15 sind die deskriptiven Statistiken der Beurteilung der Instruktionsverständlichkeit durch die MWB zu entnehmen. Danach haben 91.70 % der MWB 1 und 94.40 % der MWB 2 die Verständlichkeit der spezifischen Strategie mit einem Wert von mindestens 10 eingestuft. Für die statistische Überprüfung der Gleichheitshypothese, sprich gleiche Verständlichkeit der Instruktion für MWB 1 und 2, wurde die Signifikanzschranke auf  $p = .20$  gesetzt.<sup>54</sup> Der Vorzeichenrangtest von Wilcoxon ergab aufgrund des strengen Vorgehens einen signifikanten Unterschied ( $z = -1.360$ ,  $p = .174$ ,  $r = -.160$ ,  $1 - \beta = .31$ ; s. Abbildung 14).

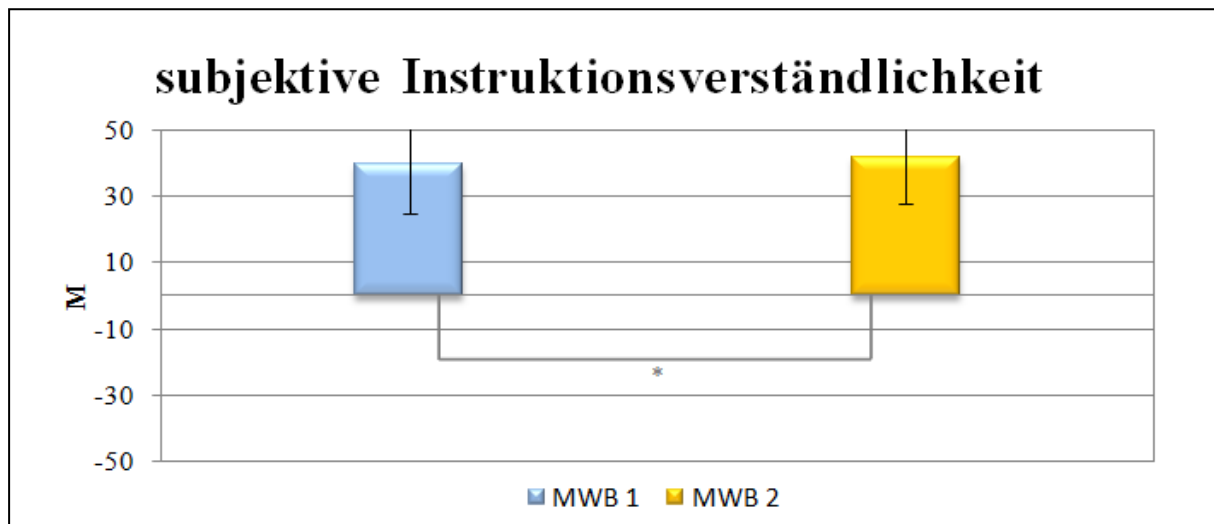


Abbildung 14: Mittelwert (M) und Standardabweichung (SD) der subjektiven Instruktionsverständlichkeit von MWB 1 und 2

Anmerkungen. Skala: -50 = gar nicht verständlich bis 50 = völlig verständlich; signifikante Unterschiede sind markiert: \* =  $p_{\text{zweiseitig}} < .20$

## 1.2 Kontrollhypothese bezüglich Instruktionsbefolgung

Die *Hypothese K.1.2* formuliert, dass MWB 1 und 2 angeben, die jeweilig instruierte spezifische Strategie befolgt zu haben und diesbzgl. kein Unterschied zwischen ihnen besteht.

Tabelle 15 sind die deskriptiven Statistiken der Beurteilung der Instruktionsbefolgung durch die MWB zu entnehmen. Danach haben 88.90 % der MWB 1 und 80.60 % der MWB 2 die Befolgung der spezifischen Strategie mit einem Wert von mindestens 10 eingestuft (Skala: -50 = gar nicht befolgt bis 50 = vollends befolgt). Zur statistischen Überprüfung der Hypothese kam abermals der Vorzeichenrangtest von Wilcoxon zum Einsatz. Zwischen den beiden MWBn liegt kein signifikanter Unterschied hinsichtlich der Instruktionsbefolgung vor ( $z = -0.126$ ,  $p = .900$ ,  $r = -.015$ ,  $1 - \beta = .20$ ).

<sup>54</sup> Dadurch wird die Wahrscheinlichkeit einer Fehlentscheidung zugunsten der Nullhypothese (Gleichheitshypothese) minimiert (Bortz, 1999).

### 1.3 Kontrollhypothese bezüglich der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen

Die *Hypothese K.1.3* postuliert, dass es MWB 1 und 2 subjektiv wichtig ist, ein gutes Tracking-Ergebnis zu erreichen. Zwischen den beiden MWBn besteht kein Unterschied hinsichtlich dieser Selbsteinschätzung. In Tabelle 15 befinden sich die deskriptiven Statistiken zur Selbstbeurteilung der MWB. Demzufolge bewerteten nur 8.30 % der MWB 1 und keiner der MWB 2 die Wichtigkeit mit Werten kleiner -10 (Skala: -50 = unwichtig bis 50 = sehr wichtig). Der Vorzeichenrangtest von Wilcoxon zur statistischen Überprüfung dieser Hypothese ergab keinen signifikanten Unterschied zwischen den beiden MWBn bzgl. der Wichtigkeit, ein gutes Tracking-Ergebnis zu erreichen ( $z = -0.344$ ,  $p = .731$ ,  $r = -.041$ ,  $1 - \beta = .28$ ).

### 1.4 Kontrollhypothesen bezüglich worauf beim Tracking geachtet wurde

*Hypothese K.1.4.1* postuliert für das Selbsturteil von MWB 1 und 2 hinsichtlich der Frage „Worauf wurde beim Tracking mehr geachtet: Geschwindigkeit vs. Genauigkeit?“ einen Unterschied zwischen ihnen. Dabei wird angenommen, dass aufgrund der jeweiligen spezifischen Strategie MWB 1 mehr auf die Geschwindigkeit und MWB 2 mehr auf die Genauigkeit geachtet hat. Zur Überprüfung dieser Annahme wurde eine zweifaktorielle univariate Varianzanalyse mit Messwiederholung durchgeführt. Neben dem Messwiederholungsfaktor „MWB“ (2fach gestuft) wurde der Messwiederholungsfaktor „Messzeitpunkt“ (4fach gestuft: MD8, 12, 16, Post)<sup>55</sup> in die statistische Analyse mit aufgenommen. Dieses Vorgehen dient u. a.<sup>56</sup> der Überprüfung der *Hypothese K.1.4.2*, wie zeitlich stabil die Präferenz der MWB ist, das Objekt schnell oder genau zu regeln.

Für den Messwiederholungsfaktor „*MWB*“ zeigt sich ein höchst signifikanter Einfluss, worauf beim Tracking geachtet wurde ( $F[1, 35] = 35.337$ ,  $p_{\text{zweiseitig}} < .001$ ,  $\eta^2_p = .500$ ,  $1 - \beta > .999$ ,  $p_{\text{einseitig}} < .001$ ; s. Tabelle 16 und Abbildung 15). Daher wurde für jeden der vier Messzeitpunkte ein Einzelvergleich durchgeführt. Aufgrund der damit verbundenen Alpha-Fehler-Kumulierung wurde die Signifikanzschranke nach der Bonferoni-Alpha-Fehler-Korrektur auf  $p = .0125$  gesetzt. Wie sich zeigt, gilt der signifikante Unterschied zwischen MWB 1 und 2 ausnahmslos für jeden Messzeitpunkt (*MD8*:  $z = -3.448$ ,  $p_{\text{einseitig}} < .001$ ,  $r = -.406$ ,  $1 - \beta = .94$ ; *MD12*:  $z = -3.991$ ,  $p_{\text{einseitig}} < .001$ ,  $r = -.470$ ,  $1 - \beta > .999$ ; *MD16*:  $z = -4.595$ ,  $p_{\text{einseitig}} < .001$ ,  $r = -.542$ ,  $1 - \beta > .999$ ; *MPost*:  $z = -4.541$ ,  $p_{\text{einseitig}} < .001$ ,  $r = -.535$ ,  $1 - \beta > .999$ ).

---

<sup>55</sup> Zum ersten Messzeitpunkt (Prä1; s. Tabelle 10) nach dem ersten Probedurchgang war den MWBn deren spezifische Strategie noch nicht bekannt. Deshalb gehen die Daten dieser Erhebung nicht in diese Auswertung ein.

<sup>56</sup> Zudem wird damit die Teststärke für die Überprüfung des Faktors „MWB“ erhöht (s. Sedlmeier & Renkewitz, 2008).

Tabelle 16: Mittelwert und Standardabweichung der Selbsturteile der MWB worauf beim Tracking mehr geachtet wurde je Messzeitpunkt und über alle Messzeitpunkte hinweg

	MWB 1		MWB 2		MWB Gesamt	
	M	SD	M	SD	M	SD
Präferenz MD8	-0.39	22.05	-20.61	20.07	-10.50	23.28
Präferenz MD12	5.64	17.79	-17.08	22.02	-5.72	22.93
Präferenz MD16	8.56	14.83	-17.36	21.09	-4.40	22.66
Präferenz MPost	8.08	16.94	-17.57	20.80	-4.74	22.83
Präferenz MGesamt	5.47	18.25	-18.15	21.02		

Anmerkungen. Skala: -50 = Genauigkeit bis 50 = Geschwindigkeit; M = Mittelwert; SD = Standardabweichung; MD = Messzeitpunkt nach Durchgang; MPost = Messzeitpunkt nach allen Durchgängen; MGesamt = über alle Messzeitpunkte hinweg

Für den Faktor „**Messzeitpunkt**“ ergibt sich ebenfalls ein sehr signifikanter Unterschied worauf beim Tracking geachtet wurde ( $F[2.801, 98.030]^{57} = 6.280, p = .001, \eta^2_p = .152, 1 - \beta > .999$ ). Zur Ermittlung, ob dieses Ergebnis sowohl durch MWB 1 als auch 2 verursacht wurde, wurde je MWB eine einfaktorielle Varianzanalyse mit dem Messwiederholungsfaktor „Messzeitpunkt“ (4fach gestuft) durchgeführt.

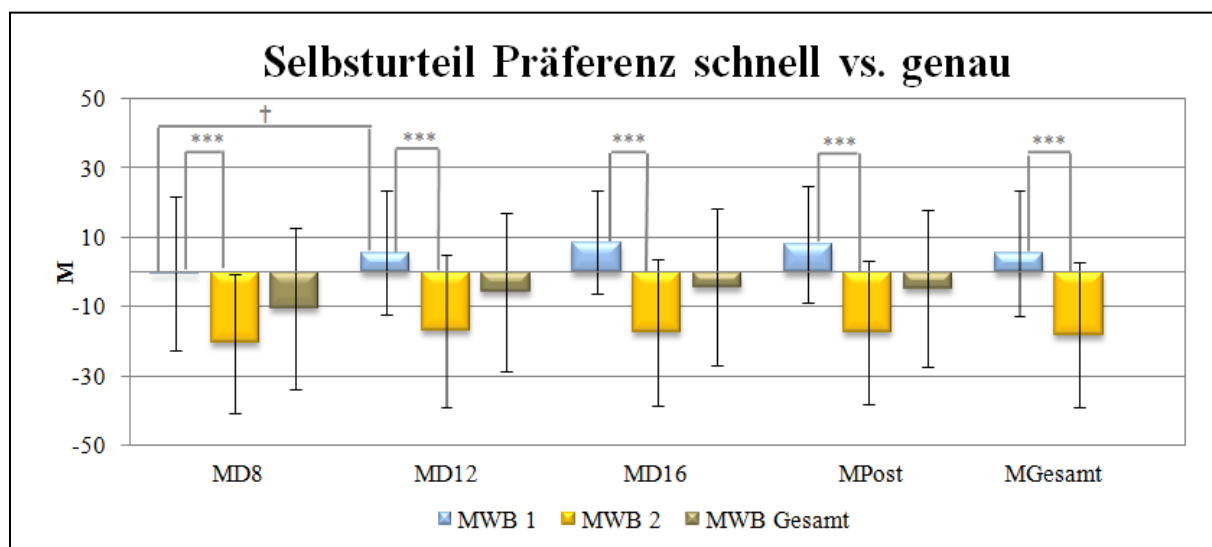


Abbildung 15: Mittelwert (M) und Standardabweichung des Selbsturteils der MWB worauf beim Tracking mehr geachtet wurde je Messzeitpunkt und über alle Messzeitpunkte hinweg

Anmerkungen. Skala: -50 = Genauigkeit bis 50 = Geschwindigkeit; MWB 1 hat spezifische Strategie schnell zu agieren, MWB 2 hat spezifische Strategie genau zu agieren; signifikante Unterschiede sind markiert: \*\*\* =  $p_{\text{einseitig}} < .001$  respektive die Alpha-Fehler-Korrektur dieser Signifikanzschranke und † =  $p_{\text{zweiseitig}} < .033$  (Signifikanzschranke aufgrund Alpha-Fehler-Korrektur)

Beim **MWB 1** liegt für den Messwiederholungsfaktor „**Messzeitpunkt**“ ebenfalls ein sehr signifikanter Unterschied vor bzgl. worauf beim Tracking geachtet wurde ( $F[2.545,$

<sup>57</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Huynh-Feldt Verfahrens korrigiert.

89.061]<sup>58</sup> = 6.882,  $p = .001$ ,  $\eta^2_p = .164$ ,  $1 - \beta > .999$ ). Zur Überprüfung wann es diesbzgl. zu Veränderungen im Verlauf der Untersuchung kam, wurden nur Einzelvergleiche zwischen den aufeinanderfolgenden Messzeitpunkten durchgeführt. Die Unterschiede zwischen diesen spiegeln die zeitlichen Veränderungen der Präferenz, das Objekt schnell zu regeln, am deutlichsten wider. Auf die weiteren möglichen Einzelvergleiche wurde verzichtet, um die Alpha-Fehler-Kumulierung möglichst klein zu halten. Aufgrund der drei ausgeführten Einzelvergleiche wurde erneut eine Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .033$ <sup>59</sup> angewendet. Bei den drei Einzelvergleichen mittels des t-Tests für abhängige Stichproben zeigt sich nur für „**MD8 vs. MD12**“ ein signifikantes Ergebnis („MD8 vs. MD12“:  $SE_{MD8} = 3.67$ ,  $SE_{MD12} = 2.97$ ,  $t[35] = -2.311$ ,  $p = .027$ ,  $r = .364$ ,  $1 - \beta = 0.55$ ; „MD12 vs. MD16“:  $SE_{MD12} = 2.97$ ,  $SE_{MD16} = 2.47$ ,  $t[35] = -1.504$ ,  $p = .142$ ,  $r = .246$ ,  $1 - \beta = .25$ ; „MD16 vs. MPost“:  $SE_{MD16} = 2.47$ ,  $SE_{MPost} = 2.82$ ,  $t[35] = -0.228$ ,  $p = .821$ ,  $r = .039$ ,  $1 - \beta = .04$ ).

Beim MWB 2 ergibt sich für den Faktor „Messzeitpunkt“ kein signifikanter Unterschied ( $F[2.457, 85.987]$ <sup>60</sup> = 0.985,  $p = .392$ ,  $\eta^2_p = .027$ ,  $1 - \beta = .70$ ).

### 3.2.3.3.4 Ergebnisse der Analysen zur Kontrollfragestellung 2

Die zweite Kontrollfragestellung geht der Entstehung eines Konflikts zwischen den MWBn nach, indem diese eine Leistungsdifferenz zwischen sich wahrnehmen. Nachfolgend werden die Ergebnisse der statistischen Überprüfungen der aus dieser Fragestellung abgeleiteten Hypothesen präsentiert.

#### 2.1 Kontrollhypothesen bezüglich worauf beim Tracking geachtet wurde

Aufgrund der konträren Ausrichtung der spezifischen Strategie von MWB 1 und 2 werden hinsichtlich deren Selbstbeurteilung und deren Beurteilung des anderen MWBs, worauf beim Tracking mehr geachtet wurde, Unterschiede angenommen. Daher wird in der **Hypothese K.2.1.1** formuliert, dass MWB 1 angibt, selbst auf die Geschwindigkeit geachtet zu haben. Wohingegen dieser einschätzt, MWB 2 habe auf die Genauigkeit geachtet. Zugleich wird in der **Hypothese K.2.1.2** postuliert, dass MWB 2 angibt, selbst auf die Genauigkeit geachtet zu haben. Wohingegen dieser einschätzt, MWB 1 habe eher auf die Geschwindigkeit geachtet. Zur Über-

---

<sup>58</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Huynh-Feldt Verfahrens korrigiert.

<sup>59</sup> Aufgrund der drei vorgenommenen Einzelvergleiche wird die Signifikanzschranke zur Vermeidung einer Alpha-Fehler-Kumulierung von  $p = .10$  auf  $.033$  gesetzt. Zuvor war bereits der Alpha-Fehler-Kumulierung aufgrund der Berechnung einer einfaktoriellen univariaten Varianzanalyse mit Messwiederholung für jeden MWB einzeln mit einer Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .10$  entgegengewirkt worden.

<sup>60</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Huynh-Feldt Verfahrens korrigiert.

prüfung dieser Hypothesen dient der Vorzeichenrangtest von Wilcoxon. Zwischen der **Selbstbeurteilung von MWB 1** und der **Fremdbeurteilung des MWBs 2** zeigt sich ein höchst signifikanter Unterschied ( $z = -3.567$ ,  $p_{\text{zweiseitig}} < .001$ ,  $r = -.420$ ,  $1 - \beta > .999$ ,  $p_{\text{einseitig}} < .001$ ; s. Abbildung 16). Ebenso findet sich zwischen der **Selbstbeurteilung von MWB 2** und der **Fremdbeurteilung des MWBs 1** ein sehr signifikanter Unterschied ( $z = -2.668$ ,  $p_{\text{zweiseitig}} = .008$ ,  $r = -.314$ ,  $1 - \beta = .96$ ,  $p_{\text{einseitig}} = .004$ ).

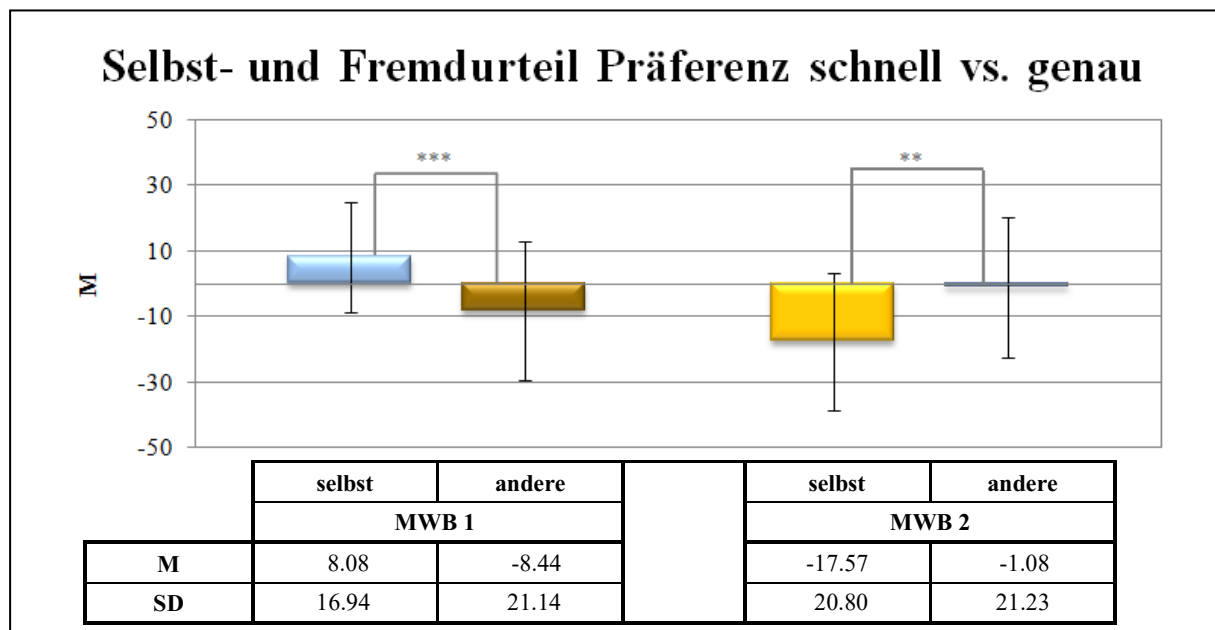


Abbildung 16: Mittelwerte (M) und Standardabweichungen (SD) der Beurteilungen der eigenen Präferenz und der des anderen MWBs hinsichtlich einer schnellen oder genauen Objektregelung

Anmerkungen. Skala: -50 = Genauigkeit bis 50 = Geschwindigkeit; signifikante Unterschiede sind markiert: \*\* =  $p_{\text{einseitig}} < .01$  und \*\*\* =  $p_{\text{einseitig}} < .001$

## 2.2 Kontrollhypothese bezüglich der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen

In der **Hypothese K.2.2** wird zwischen der Selbstbeurteilung und der Beurteilung des anderen MWBs, wie wichtig es war ein gutes Tracking-Ergebnis zu erreichen, ein Unterschied formuliert. Die MWB beurteilen die Wichtigkeit, ein gutes Tracking-Ergebnis zu erreichen jeweils für sich selbst höher als für den anderen MWB. Zur Überprüfung dieser Annahme wurde eine zweifaktorielle univariate Varianzanalyse mit Messwiederholung durchgeführt.

Tabelle 17: Mittelwerte und Standardabweichungen der Beurteilungen der eigenen Wichtigkeit und der des anderen MWBs ein gutes Tracking-Ergebnis zu erreichen

	MWB 1		MWB 2		MWB Gesamt	
	M	SD	M	SD	M	SD
selbst	26.03	18.74	28.58	15.98	27.31	17.34
andere	28.19	14.35	26.50	17.70	27.35	16.02
selbst/andere Gesamt	27.11	16.61	27.54	16.78		

Anmerkungen. Skala: -50 = unwichtig bis 50 = sehr wichtig; M = Mittelwert; SD = Standardabweichung

Für den Messwiederholungsfaktor „MWB selbst vs. anderer“ (2fach gestuft) ergibt sich kein signifikanter Unterschied hinsichtlich der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen ( $F[1, 35] = 0.002$ ,  $p_{\text{zweiseitig}} = .969$ ,  $\eta^2_p < .001$ ,  $1 - \beta = .05$ ,  $p_{\text{einseitig}} = .485$ ; s. Tabelle 17). Für den Messwiederholungsfaktor „MWB“ (2fach gestuft) zeigt sich ebenfalls kein signifikanter Einfluss auf die Beurteilung der Wichtigkeit, ein gutes Tracking-Ergebnis zu erreichen ( $F[1, 35] = 0.015$ ,  $p = .902$ ,  $\eta^2_p < .001$ ,  $1 - \beta = .20$ ).

### 2.3 Kontrollhypothese bezüglich der Zufriedenheit

Die **Hypothese K.2.3** postuliert, dass bei MWB 1 und 2 die Zufriedenheit mit der Teamleistung nach den kooperativen Tracking-Durchgängen eher mittelmäßig bis gering ausfällt und diesbzgl. kein Unterschied zwischen ihnen besteht. Tabelle 18 sind die deskriptiven Statistiken der Beurteilung der Zufriedenheit durch die MWB zu entnehmen. Demzufolge beurteilen nur 5.60 % der MWB 1 und 2 die Zufriedenheit mit Werten kleiner -10. Zur statistischen Überprüfung dieser Hypothese kam erneut der Vorzeichenrangtest von Wilcoxon zur Anwendung. Es liegt kein signifikanter Unterschied zwischen MWB 1 und 2 hinsichtlich der Zufriedenheit mit der Teamleistung nach den kooperativen Durchgängen vor ( $z = -0.491$ ,  $p = .623$ ,  $r = -.058$ ,  $1 - \beta = .22$ ). Die mittlere Zufriedenheit der MWB 1 und 2 liegt im positiven Bereich, anders als intendiert.

Tabelle 18: Deskriptive Statistiken der Beurteilung der Zufriedenheit mit der Teamleistung nach den kooperativen Tracking-Durchgängen durch die MWB

	M	SD	Median	% der MWB mit SW > 10	% der MWB mit SW < -10
MWB 1	23.14	17.96	24.50	83.30	5.60
MWB 2	24.42	21.15	27.50	75.00	5.60

Anmerkungen. Skala: -50 = sehr unzufrieden mit dem Team bis 50 = sehr zufrieden mit dem Team; M = Mittelwert; SD = Standardabweichung; % = Prozent; SW = Skalenwert

### 2.4 Kontrollhypothesen bezüglich der Stimmung

Die **Hypothese K.2.4.1** postuliert, dass zwischen den beiden MWBn kein Unterschied bei der Beurteilung der eigenen Stimmung existiert. In der **Hypothese K.2.4.2** wird formuliert, dass die Stimmung der beiden MWB vom Messzeitpunkt vor (MPrä2) allen messrelevanten Tracking-Durchgängen zum Messzeitpunkt danach (MPost) abnimmt. Zur Überprüfung dieser beiden Annahmen wurde eine zweifaktorielle univariate Varianzanalyse mit Messwiederholung durchgeführt. Für den Messwiederholungsfaktor „MWB“ (2fach gestuft) ergibt sich kein signifikanter Unterschied bei der Selbstbeurteilung der Stimmung ( $F[1, 35] = 1.511$ ,  $p = .227$ ,  $\eta^2_p = .041$ ,  $1 - \beta = .99$ ; s. Abbildung 17). Für den Messwiederholungsfaktor „Messzeitpunkt“ (2fach gestuft: MPrä2 vs. MPost) besteht ein signifikanter Unterschied hin-

sichtlich der Stimmung ( $F[1, 35] = 8.549$ ,  $p_{\text{zweiseitig}} = .006$ ,  $\eta^2_p = .196$ ,  $1 - \beta > .999$ ,  $p_{\text{einseitig}} = .003$ ). Trotzdem liegt die durchschnittliche Stimmung der MWB auch nach den Tracking-Durchgängen immer noch im positiven Bereich, anders als intendiert.

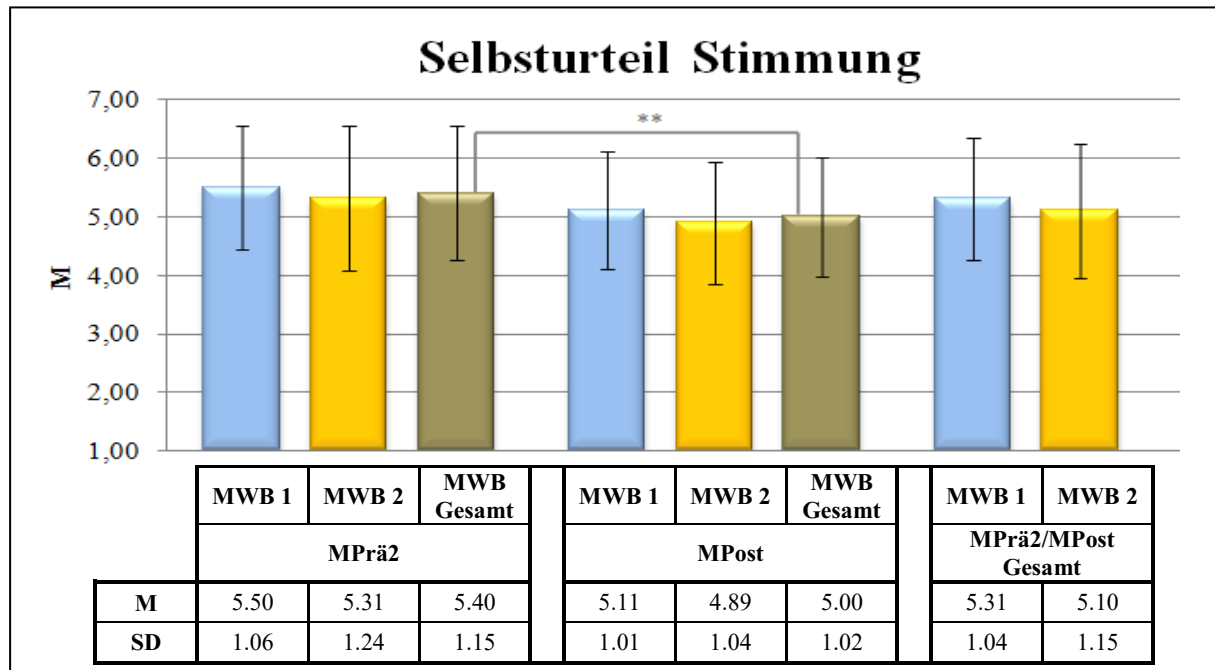


Abbildung 17: Mittelwert (M) und Standardabweichung (SD) des Selbsturteils der Stimmung der beiden MWB von den Messzeitpunkten vor (MPrä2) und nach (MPost) den messrelevanten kooperativen Tracking-Durchgängen sowie über diese Messzeitpunkte hinweg

*Anmerkungen.* Kunin-Skala: 1 = negativstes Emoticon bis 7 = positivstes Emoticon; signifikante Unterschiede sind markiert: \*\* =  $p_{\text{einseitig}} < .01$

## 2.5 Kontrollhypothese bezüglich Gleichheit der instruierten Tracking-Strategien

Die **Hypothese K.2.5** postuliert, dass bei allen MWB-Teams beide MWB die gesamte Untersuchung davon ausgehen, dieselbe spezifische Strategie für das Tracking instruiert bekommen zu haben. Zur Überprüfung dieser Annahme wurde die Anzahl der MWB-Paare, in denen die Unterschiedlichkeit der beiden instruierten spezifischen Strategien durch mindestens einen MWB erkannt wurde, mit der Anzahl an Teams verglichen, in denen diese nicht erkannt wurde. Dafür kam der Binomialtest<sup>61</sup> zum Einsatz. Die gefundenen Anteile der MWB-Paare, welche die Unterschiedlichkeit der spezifischen Strategien erkannt vs. nicht erkannt haben, weichen höchst signifikant von den erwarteten Anteilen<sup>62</sup> ab ( $H_{\text{erkannt}} = 16$ ,  $H_{\text{nichterkannt}} = 20$ ,  $p < .001$ ,  $g = .443$ , odds ratio = 797.76,  $1 - \beta > .999$ ).

<sup>61</sup> Es wurde der Binomialtest verwendet, weil der Chi-Quadrat-Test bei einer dichotomen Variable je Kategorie eine erwartete Häufigkeiten von mindestens zehn voraussetzt (s. Bortz & Lienert, 1998).

<sup>62</sup> Die Autorin erwartete einen Anteil von null MWB-Teams, bei denen mindestens ein MWB die Unterschiedlichkeit der spezifischen Strategien erkannt hat. Daher wurde der Binomialtest mit einem Testanteil von .001 durchgeführt.



### 3.2.3.4 Interpretation und Diskussion der Ergebnisse

In diesem Kapitel erfolgt eine zusammenfassende Darstellung, Interpretation und Diskussion der gefundenen Ergebnisse.

#### 3.2.3.4.1 Interpretation und Diskussion der Ergebnisse zur Forschungsfragestellung 3

Hinsichtlich der durch die MWB verursachten Anzahl an Fehlern beim Tracking finden sich zwischen SAM 1.0 und 2.0 bei allen sechs betrachteten Durchgängen höchst signifikante hypothesenkonforme Unterschiede mit extrem hohen Teststärken. Diese Unterschiede sind zudem praktisch bedeutsam. Die gerichtete *Hypothese F.3.1*, dass unter Einsatz von SAM 2.0 die MWB mehr Fehler produzieren als bei SAM 1.0, kann **angenommen** werden.

Wie die inferenzstatistische Analyse zeigte, benötigen die MWB hypothesenkonform im Durchschnitt unter Einsatz von SAM 2.0 über alle betrachteten Durchgänge hinweg signifikant weniger Zeit zum Tracking als bei SAM 1.0. Auf Ebene der einzelnen Durchgänge liegt zwischen den beiden SAM-Versionen diesbzgl. lediglich bei drei Durchgängen ein tendenziell signifikanter Unterschied vor. Trotzdem weisen alle Durchgänge mittelgroße Unterschiede zwischen den beiden SAM-Versionen auf. Aufgrund der geringen Teststärken lag die Wahrscheinlichkeit, ein signifikantes Ergebnis zu finden, jedoch nur bei ca. 40 %. Eine Ausnahme stellt Durchgang 16 dar. Bei diesem besteht nur ein kleiner bis mittlerer Unterschied zwischen den SAM-Versionen sowie eine extrem kleine Teststärke. Dies kann durch das Ausreißer-MWB-Team der Stichprobe von SAM 1.0 verursacht sein, welches eine extrem geringe Zeit in Durchgang 16 benötigte. Bei Ausschluss der Ausreißer aus den Analysen existiert sogar bei Durchgang 16 ein annähernd mittelgroßer Unterschied. Mittlere Effektgrößen werden von der Autorin als praktisch bedeutsam erachtet. Dennoch bleibt die Teststärke mit  $1 - \beta = .57$  zu klein für eine statistische Absicherung. Insgesamt liegt nun für den Großteil der Durchgänge eine hypothesenkonforme Signifikanz vor. Da jedoch nicht bei jedem Durchgang ein signifikanter Unterschied zwischen den beiden SAM-Versionen besteht, kann die *Hypothese F.3.2*, dass unter Einsatz von SAM 2.0 weniger Zeit zum Tracking benötigt wird als bei SAM 1.0, **nicht angenommen** werden. Trotzdem stimmen die gefundenen Ergebnisse optimistisch, dass die vorgenommenen Veränderungen, wie bspw. die Geschwindigkeitsanzeige, in SAM 2.0 die MWB dazu bewegen, das Tracking-Objekt mit einer höheren Geschwindigkeit zu regeln.

Zudem liegt bei beiden SAM-Versionen eine signifikante Verringerung der benötigten Zeit zum Tracking über die Untersuchung hinweg vor. Wohingegen die Anzahl der Fehler beim Tracking über die Zeit stabil blieb. Hierin könnte sich ein Lerneffekt der MWB widerspiegeln.

Unterstützt wird diese Annahme dadurch, dass die MWB von SAM 2.0<sup>63</sup>, trotz der schnelleren Objektregelung, die einzelnen Durchgänge als gleich beanspruchend ( $F[3.311, 115.878]^{64} = 0.409$ ,  $p = .766$ ,  $\eta^2_p = .012$ ,  $1 - \beta = .56$ ) und komplex ( $F[3.049, 106.710]^{65} = 1.875$ ,  $p = .137$ ,  $\eta^2_p = .051$ ,  $1 - \beta > .999$ ) empfanden (s. Tabelle 19).

Tabelle 19: Mittelwert und Standardabweichung der subjektiven Beanspruchung der MWB und subjektiven Komplexität der einzelnen Tracking-Durchgänge für beide MWB zusammen pro Messzeitpunkt

	MD7	MD8	MD10	MD12	MD14	MD16
<b>subjektive Beanspruchung</b>						
<b>M</b>	55.08	54.15	56.07	55.54	55.93	56.15
<b>SD</b>	20.93	22.65	21.13	21.23	20.71	20.99
<b>subjektive Komplexität</b>						
<b>M</b>	-5.53	-7.26	-3.28	-4.51	-6.10	-4.72
<b>SD</b>	16.94	17.25	15.11	15.86	15.87	15.25

Anmerkungen. Skala „Beanspruchung“: 0 = absolut keine Anstrengung bis 150 = unerträgliche Anstrengung; Skala „Komplexität“: -50 = niedrige Komplexität bis 50 = hohe Komplexität; MD = Messzeitpunkt nach Durchgang x; M = Mittelwert; SD = Standardabweichung

Außerdem fühlten sich die MWB im Laufe der Untersuchung *sicherer* im *Umgang* mit dem *Joystick* ( $M_{MP\text{Prä}2} = 02.47^{66}$ ,  $SD_{MP\text{Prä}2} = 22.52$ ,  $M_{MP\text{Post}} = 09.74$ ,  $SD_{MP\text{Post}} = 20.40$ ,  $F[1, 35] = 22.500$ ,  $p < .001$ ,  $\eta^2_p = .391$ ,  $1 - \beta > .999$ ). Scheinbar hatten die MWB die Regelung der Objektgeschwindigkeit an deren steigende Fertigkeit im Umgang mit dem Joystick so angepasst, dass das Niveau der Objektgenauigkeit gleichblieb.

Dieser Lerneffekt beim kooperativen Tracking könnte zudem einen Anteil an der unterschiedlichen durchschnittlich benötigten Zeit zwischen den beiden SAM-Versionen haben. In SAM 2.0 konnten die MWB bereits im Durchgang 5 und 6 erste Erfahrungen mit dem kooperativen Tracking sammeln. Somit war es diesen MWBn möglich, bereits in diesen beiden Durchgängen eine Fertigkeit im kooperativen Tracking aufzubauen. Wohingegen in SAM 1.0 das erste kooperative Tracking erst im Durchgang 7 stattfindet. Zudem hatten die MWB von SAM 2.0 vier weitere Durchgänge (9, 11, 13, 15) zum Üben des kooperativen Trackings, welche in SAM 1.0 noch solo absolviert werden müssen.

Dieser Lerneffekt sowie die unterschiedliche Anzahl an kooperativen Tracking-Durchgängen der beiden SAM-Versionen können ferner als Erklärung herangezogen werden,

<sup>63</sup> Von der Stichprobe der SAM-Version 1.0 liegen der Autorin keine Fragebogendaten vor, so dass zur mentalen Beanspruchung dieser MWB sowie über deren Wahrnehmung der Komplexität der Tracking-Durchgänge keine Aussagen getroffen werden können.

<sup>64</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Greenhouse-Geisser Verfahrens korrigiert.

<sup>65</sup> Aufgrund verletzter Sphärizitätsannahme wurden die Freiheitsgrade mittels des Greenhouse-Geissers Verfahren korrigiert.

<sup>66</sup> Skala: - 50 = ich bin sehr unsicher wie das geht bis 50 = ich bin sehr sicher wie das geht

warum beim Durchgang 16 zwischen den Versionen der kleinste Unterschied hinsichtlich der benötigten Zeit zum Tracking vorliegt. Es ist anzunehmen, dass die Lernkurve der MWB bzgl. der Objektregelung in beiden SAM-Versionen aufgrund der Ausbildung der gleichen motorischen Fertigkeit annähernd gleich verlief. Durch die höhere Anzahl an kooperativen Tracking-Durchgängen in SAM 2.0 lagen diese MWB denen von SAM 1.0 beim Lernen voraus. Möglicherweise hatten die MWB von SAM 2.0 beim Durchgang 16 ein Lernplateau erreicht oder es lag bereits die Annäherung an den Lerngrenzwert vor, wodurch nur noch geringere Lernfortschritte zu erreichen waren (s. Loosch, 1999). Dadurch wäre es den MWBn von SAM 1.0 hinsichtlich der benötigten Zeit beim Tracking möglich näher an die MWB von SAM 2.0 heranzukommen. In SAM existiert ein begrenzter Raum, in dem sich die MWB verbessern können. Die Geschwindigkeitsstufen, mit denen das Tracking-Objekt geregelt werden kann, entsprechen der Joystick-Auslenkung auf der Vertikalen.

Des Weiteren zeigen sich bei detaillierter Betrachtung der objektiven Verhaltensdaten, wie in den Voruntersuchungen SAM 1.5 (s. Anhang A.3.3), Unterschiede zwischen den MWBn 1 und 2. Die MWB 1 führten im Durchschnitt schnellere und stärkere positive Joystick-Auslenkungen auf der Vertikalen aus. Wohingegen die MWB 2 im Mittel stärkere negative Joystick-Auslenkungen auf der Vertikalen vornahmen.

#### **3.2.3.4.2 Interpretation und Diskussion der Ergebnisse zur Forschungsfragestellung 4**

Die vergleichenden inferenzstatistischen Analysen zwischen dem kurzen und dem langen Tracking-Pfad, um deren Eignung für SAM zu ermitteln, erbrachten bei den meisten  $aV_n$  Unterschiede. Hierbei ist zu beachten, dass sich die verglichenen Werte der  $aV_n$  „benötigte Zeit“ und „Flächenfehler“ beim Tracking auf die gleiche Pfadlänge beziehen (s. Abschnitt 3.2.1.3.1).

Während bei dem kurzen Tracking-Pfad im Mittel „eher viel Anstrengung“ von den MWBn wahrgenommen wurde, nähert sich deren empfundene Beanspruchung durch den langen Tracking-Pfad bereits an den verbalen Anker „erhebliche Anstrengung“ an. Dieser Unterschied ist höchst signifikant und weist eine sehr hohe praktische Bedeutsamkeit auf. Ebenso wurde der lange Tracking-Pfad von den MWBn im Durchschnitt höchst signifikant und praktisch bedeutsam als komplexer wahrgenommen. Trotz der höheren Beanspruchung der MWB und Komplexität der Tracking-Aufgabe durch den längeren Tracking-Pfad liegt für die  $aV$  „benötigte Zeit“ zum Tracking bei diesem ein signifikant geringerer Wert vor. Dieser Unterschied weist jedoch bei Ausschluss der zwei Ausreißer nur noch eine annähernd mittlere praktische Bedeutsamkeit auf, welche aufgrund der gleichzeitig gesunkenen Teststärke statistisch nicht mehr abgesichert werden kann. Hinsichtlich der  $aV$  „Flächenfehler“ beim Tracking

findet sich kein signifikanter Unterschied. Es ist anzunehmen, dass die Fertigkeit der MWB zum Tracking weiterhin gestiegen war wie in den Durchgängen zuvor. Die MWB scheinen deren Objektregelungsverhalten an ein bestimmtes Genauigkeitsniveau angepasst zu haben. Die Objektgeschwindigkeit dagegen variiert je nach Genauigkeitsniveau, Tracking-Fertigkeit und Komplexität der Tracking-Aufgabe. Das unterschiedliche Genauigkeitsniveau zwischen den beiden SAM-Versionen kann wiederum durch die unterschiedliche Motivation der MWB verursacht sein, welche in SAM 2.0 höher ausfallen sollte.<sup>67</sup>

Wenn auch hinsichtlich der Fehleranzahl und der benötigten Zeit zum Tracking unter Ausschluss der Ausreißer keine praktisch bedeutsamen Unterschiede zwischen dem kurzen und dem langen Tracking-Pfad vorliegen, so scheint dennoch der lange Tracking-Pfad geeigneter für die Verwendung in SAM zu sein. Die subjektive Beanspruchung der MWB beim Tracking kann als ein Maß für die Güte der Regelungstätigkeit von Operateuren und Automaten verwendet werden. Je höher die subjektive Komplexität der Tracking-Aufgabe ist und dadurch bedingt die subjektive Beanspruchung der MWB ohne Einsatz von Operateuren und Automaten, desto mehr Optimierungspotential von SAM besteht für die Operateure und Automaten. Gleichzeitig bietet der längere Tracking-Pfad den Operateuren und Automaten einen längeren Einsatz sowie damit einhergehend mehr Eingriffsmöglichkeiten.

#### 3.2.3.4.3 Interpretation und Diskussion der Ergebnisse zur Kontrollfragestellung 1

Zwischen den MWBn 1 und 2 liegt, anders als bei den Voruntersuchungen SAM 1.5 (s. Anhang B.4.3), bzgl. der subjektiven Verständlichkeit der jeweiligen spezifischen Strategie ein signifikanter Unterschied vor. Die MWB 2 fanden ihre Genauigkeitsstrategie verständlicher, als die MWB 1 ihre Geschwindigkeitsstrategie. Die *Hypothese K.1.1*, dass zwischen dem MWB 1 und 2 kein Unterschied hinsichtlich der Verständlichkeit der jeweiligen spezifischen Strategie besteht, muss daher **verworfen** werden. Jedoch fällt der gefundene Unterschied zwischen den MWBn 1 und 2 eher klein aus, so dass dieser von der Autorin als praktisch bedeutungslos eingestuft wird.

Zur Ermittlung, ob die beiden spezifischen Strategien von den MWBn als ausreichend verständlich empfunden wurden, wurde ein Vorgehen aus der ersten ATEO-Projektphase angewendet (s. Nachtwei, 2008a). Danach dürfen zum einen nicht mehr als 10 % der MWB einen Wert kleiner -10 (entspricht wenig bis gar nicht verständlich) vergeben haben. Zum anderen müssen ca. 66 % der MWB mindestens einen Wert von 10 (entspricht ziemlich bis hin

---

<sup>67</sup> Weil der Autorin keine Information über die Motivation der MWB von SAM 1.0 vorliegen, kann diese Annahme nicht belegt werden.

zu völlig verständlich) angegeben haben. Die Beurteilungen der Verständlichkeit der jeweiligen spezifischen Strategie durch die MWB erfüllen diese Kriterien. Die spezifischen Strategien wurden von den MWBn wie bereits in den Voruntersuchungen SAM 1.5 (s. Anhang B.4.3) im Mittel als sehr verständlich eingestuft. Wobei diese diesmal als etwas verständlicher bewertet wurden.

Überdies bestehen zwischen den MWBn 1 und 2 keine signifikanten Unterschiede hinsichtlich der subjektiven Befolgung der jeweiligen spezifischen Strategie und der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen. Ebenso fallen die Effektgrößen dieser nicht signifikanten Unterschiede extrem klein aus. Daher werden diese von der Autorin als vernachlässigbar erachtet. Außerdem zeigt sich, nach den beiden zuvor erläuterten Kriterien, bei den MWBn 1 und 2 eine ausreichende Befolgung der jeweiligen spezifischen Strategie sowie Wichtigkeit, ein gutes Tracking-Ergebnis zu erreichen. Im Durchschnitt gaben die MWB an, die spezifischen Strategien ziemlich verfolgt und ebenso ein gutes Tracking-Ergebnis für ziemlich wichtig erachtet zu haben. Daher wird die **Hypothese K.1.2 nicht verworfen**. Diese formuliert, dass MWB 1 und 2 die jeweilige spezifische Strategie subjektiv befolgt haben und diesbzgl. kein Unterschied zwischen ihnen existiert. Ebenfalls wird deshalb die **Hypothese K.1.3 nicht verworfen**. Diese postuliert, dass es MWB 1 und 2 subjektiv wichtig ist ein gutes Tracking-Ergebnis zu erreichen und diesbzgl. kein Unterschied zwischen ihnen existiert. Die Ergebnisse zu diesen beiden Hypothesen entsprechen den Befunden aus den Voruntersuchungen SAM 1.5 (s. Anhang B.4.3).

Des Weiteren hatten im Durchschnitt die MWB 1 mehr auf die Geschwindigkeit und die MWB 2 mehr auf die Genauigkeit geachtet. Diese konträren Ausrichtungen entsprechen der jeweiligen spezifischen Strategie. Dieser Unterschied ist mit extrem hoher Effektgröße und Teststärke hypothesenkonform signifikant. Somit kann die **Hypothese K.1.4.1 angenommen** werden. Diese postuliert, dass zwischen dem Selbsturteil von MWB 1 und 2, worauf beim Tracking mehr geachtet wurde, ein Unterschied existiert. Dieser Befund zeigte sich bereits in den Voruntersuchungen SAM 1.5 (s. Anhang B.4.3).

Bei dieser Objektregelungspräferenz bestehen bei den MWBn 1 zwischen dem Durchgang 8 und den späteren Messzeitpunkten signifikante Unterschiede. Worauf beim Tracking geachtet wurde, kann somit nicht als stabil über den Untersuchungszeitraum angesehen werden. Die **Hypothese K.1.4.2** von einer zeitlichen Stabilität der Präferenz zum schnellen oder genauen Objektregeln muss **verworfen** werden. Bei genauer Betrachtung der Präferenzwerte der MWB 1 von den vier Messzeitpunkten (MD8, MD12, MD16, MPost) ist eine durchschnittliche Veränderung in Richtung schneller Objektregelung ersichtlich, so wie es

durch deren spezifische Strategie intendiert war. Dieser Trend kann durch Lerneffekte der MWB 1 verursacht worden sein, welche sich aufgrund einer gesteigerten Tracking-Fertigkeit von Durchgang zu Durchgang hinsichtlich der Geschwindigkeit mehr zutrauten. Jedoch wurde der durch die spezifische Strategie anvisierte Bereich nicht erreicht, das Objekt eher schnell zu regeln (entspricht einem Skalenwert von 10 bis 30).

Die durchschnittliche Beurteilung der Objektregelungspräferenz der MWB 2 liegt bei allen vier Messzeitpunkten (MD8, MD12, MD16, MPost) in dem Bereich, der einem eher genauen Tracking entspricht (entspricht einem Skalenwert von -10 bis -30) und durch deren spezifische Strategie intendiert war.

Eine zusätzliche Betrachtung der Präferenzausprägungen vor der Instruktion der jeweiligen spezifischen Strategie zeigte auf, dass sowohl die MWB 1 als auch 2 im Mittel ein eher genaues Tracking bevorzugten. Diesbzgl. ergibt sich zwischen den MWBn 1 und 2 kein statistisch signifikanter sowie praktisch bedeutsamer Unterschied ( $M_{MWB1} = -25.83$ ,  $SD_{MWB1} = 17.86$ ,  $M_{MWB2} = -23.78$ ,  $SD_{MWB2} = 23.09$ ,  $z = -0.071$ ,  $p = .944$ ,  $r = -.008$ ,  $1 - \beta = .03$ )<sup>68</sup>. Ferner ergab ein Vergleich je MWB bzgl. *worauf beim Tracking vor und nach* der Vergabe der *spezifischen Strategie geachtet* wurde für **MWB 1** einen höchst signifikanten Unterschied ( $M_{MWB1MPrä1} = -25.83$ ,  $SD_{MWB1MPrä1} = 17.86$ ,  $M_{MWB1MD8} = -0.39$ ,  $SD_{MWB1MD8} = 22.05$ ,  $z = -4.337$ ,  $p < .001$ ,  $r = -.511$ ,  $1 - \beta > .999$ ). Insofern kann von einer Wirkung der spezifischen Strategie „Geschwindigkeitsorientierung“ in die beabsichtigte Richtung ausgegangen werden, wenn auch diesbzgl. noch Optimierungspotential existiert. Die Genauigkeitspräferenz der MWB vor der Instruktion der jeweiligen spezifischen Strategie kann ein weiterer Grund sein, warum die MWB 1 nach dieser Instruktion eine Präferenz zur schnelleren Objektregelung erst von Durchgang zu Durchgang entwickelten und trotzdem den durch die spezifische Geschwindigkeitsstrategie intendierten Bereich nicht erreichten.

Ebenso kann diese Genauigkeitspräferenz dazu beigetragen haben, dass die Instruktion der spezifischen Genauigkeitsstrategie von den MWBn als verständlicher eingeschätzt wurde als die der spezifischen Geschwindigkeitsstrategie.

Für MWB 2 zeigte der Vergleich der Objektregelungspräferenz vor und nach der Instruktion der spezifischen Strategie einen nicht signifikanten Unterschied auf ( $M_{MWB2MPrä1} = -23.78$ ,  $SD_{MWB2MPrä1} = 23.09$ ,  $SE_{MWB2MPrä1} = 3.85$ ,  $M_{MWB2MD8} = -20.61$ ,  $SD_{MWB2MD8} = 20.07$ ,  $SE_{MWB2MD8} = 3.35$ ,  $t[35] = -0.813$ ,  $p = .422$ ,  $r = .136$ ,  $1 - \beta = .05$ ). Dieses Ergebnis ist nicht

---

<sup>68</sup> Aufgrund der Alpha-Fehler-Kumulierung durch die drei explorativen Einzelvergleiche „MWB 1 MPrä1 vs. MWB 2 MPrä1“, „MWB 1 MPrä1 vs. MD8“, „MWB 2 MPrä1 vs. MD8“ liegt die Signifikanzschranke bei  $p = .017$ .

verwunderlich, da die MWB 2 im Durchschnitt bereits vor der Instruktion der spezifischen Strategie die dadurch zu bildende Verhaltenspräferenz beim Tracking aufwiesen.

Aus diesen Ergebnissen wird geschlussfolgert, dass die MWB ausreichend motiviert waren, deren jeweilige instruierte spezifische Strategie umzusetzen. Eine Konfundierung durch die Art der spezifischen Strategie hinsichtlich der Motivation der MWB liegt nicht vor.

#### 3.2.3.4.4 Interpretation und Diskussion der Ergebnisse zur Kontrollfragestellung 2

Bei den MWBn 1 und 2 liegen hypothesenkonforme signifikante Unterschiede zwischen dem Selbsturteil und der Beurteilung des anderen MWBs vor bzgl. worauf beim Tracking mehr geachtet wurde. Die MWB 1 gaben an, selbst auf die Geschwindigkeit geachtet zu haben. Wohingegen diese einschätzten, MWB 2 habe eher auf die Genauigkeit geachtet. Die MWB 2 gaben an, selbst auf die Genauigkeit geachtet zu haben. Dagegen glaubten diese, MWB 1 habe eher auf die Geschwindigkeit geachtet. Beide Ergebnisse weisen mittlere bis hohe Effektgrößen und extrem große Teststärken auf. Somit können die *Hypothesen K.2.1.1* und *K.2.1.2*, in denen diese Unterschiede zwischen dem Selbst- und dem Fremdurteil von MWB 1 und 2 postuliert sind, **angenommen** werden. Diese Hypothesen konnten ebenso bereits durch die Voruntersuchungen SAM 1.5 bestätigt werden (s. Anhang B.4.3).

Hinsichtlich der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen, liegen nur bei den MWBn 2 für sich selbst durchschnittlich höhere Beurteilungen vor als für den anderen MWB. Jedoch bestehen weder zwischen dem Selbsturteil der MWB im Vergleich zur Beurteilung des anderen MWBs noch zwischen MWB 1 und 2 signifikante, praktisch bedeutsame Unterschiede. Die *Hypothese K.2.2*, dass die MWB die Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen jeweils für sich selbst höher beurteilen als für den anderen MWB, muss **verworfen** werden. Trotzdem die MWB bei dem anderen MWB ein Tracking-Verhalten beobachteten, welches nicht der vermeintlichen spezifischen Team-Strategie entsprach, sprachen diese dem anderen MWB nicht ab, ein gutes Tracking-Ergebnis erreichen zu wollen. Sowohl für sich selbst als auch für den anderen MWB beurteilten diese im Durchschnitt, dass ein gutes Tracking-Ergebnis ziemlich wichtig ist.

Hinsichtlich der Selbstbeurteilung der Zufriedenheit mit der Teamleistung der MWB 1 und 2 nach den kooperativen Tracking-Durchgängen liegt hypothesenkonform kein signifikanter Unterschied zwischen diesen vor. Jedoch waren die MWB im Mittel ziemlich zufrieden. Somit muss die *Hypothese K.2.3*, dass bei MWB 1 und 2 die Zufriedenheit mit der Teamleistung nach den kooperativen Tracking-Durchgängen eher mittelmäßig bis gering ausfällt und diesbzgl. kein Unterschied zwischen ihnen besteht, **verworfen** werden. Dieses Ergebnis kann mit der Annahme von Kaufmann, Pornschlegel und Udris (1982) erklärt werden, dass sich

Unzufriedenheit durch Belastungen erst über einen mittleren bis längeren Zeitraum einstellt. Im Vergleich zu den Voruntersuchungen sind die Mittelwerte und Mediane allerdings etwas niedriger (s. Anhang A.3.3). Zudem hat sich der Anteil der MWB 2 etwas erhöht, welche eine hypothesenkonforme mittelmäßige bis geringe Zufriedenheit aufwiesen.

Bei der Selbstbeurteilung der Stimmung der MWB zeigt sich ebenfalls hypothesenkonform kein signifikanter Unterschied zwischen den MWBn 1 und 2. Die statistische Überprüfung weist eine extrem hohe Teststärke auf. Die **Hypothese K.2.4.1**, dass zwischen MWB 1 und 2 kein Unterschied bei der Beurteilung der eigenen Stimmung existiert, kann somit **nicht verworfen** werden. Ferner konnte diesmal eine hypothesenkonforme sehr signifikante negative Veränderung der Stimmung vom Zeitpunkt vor den Tracking-Durchgängen zum Zeitpunkt danach bei den MWBn bewirkt werden. Bei den Voruntersuchungen SAM 1.5 zeigte sich noch nicht eine solche Stimmungsveränderung (s. Anhang B.4.3). Die **Hypothese K.2.4.2**, welche diese negative Veränderung der Stimmung postuliert, kann somit **angenommen** werden. Allerdings liegt die durchschnittliche Stimmung der MWB nach den Tracking-Durchgängen weiterhin im positiven Bereich. Daher ist es fraglich, ob diese Stimmungsveränderung tatsächlich durch einen Konflikt zwischen den MWBn 1 und 2 verursacht ist. Ein Konflikt sollte sich stärker negativ auf die Stimmung der MWB auswirken. Für die Stimmungsveränderung könnte ebenso das Tracking selbst verantwortlich gewesen sein. Das Tracking wurde durch die MWB über die Untersuchung hinweg als eher hoch anstrengend empfunden (s. Abschnitt 3.2.3.4.2). Ebenso kann die im Verlauf der Untersuchung stattgefundene **Ermüdung** der MWB ( $M_{MP\ddot{u}2} = 07.26$ ,  $SD_{MP\ddot{u}2} = 22.86$ ,  $M_{MP\text{Post}} = -06.96$ ,  $SD_{MP\text{Post}} = 20.65$ ,  $F[1, 35] = 66.194$ ,  $p < .001$ ,  $\eta^2_p = 0.654$ ,  $1 - \beta > .999$ ) einen Beitrag zur negativen Stimmungsveränderung geleistet haben. Durch den zusätzlichen 17. Tracking-Pfad war SAM 2.0 länger und anstrengender für die MWB als SAM 1.5. Damit könnte erklärt werden, warum sich bei den Voruntersuchungen mit SAM 1.5 noch keine signifikante negative Stimmungsveränderung zeigte.

Des Weiteren hatte bei sechzehn MWB-Teams mindestens einer der beiden MWB die Unterschiedlichkeit der instruierten spezifischen Strategien für das Objekt-Tracking erkannt. Somit muss die **Hypothese K.2.5**, dass bei allen MWB-Teams beide MWB die gesamte Untersuchung davon ausgehen, dieselbe spezifische Strategie für das Tracking instruiert bekommen zu haben, **verworfen** werden. Es handelt sich zwar um eine Verbesserung im Vergleich zu den Voruntersuchungen SAM 1.5 wo noch bei 50 % der MWB-Teams die Unterschiedlichkeit der spezifischen Strategien erkannt wurde (s. Anhang B.4.3). Jedoch ist der Unterschied nicht groß. Die vorgenommenen Veränderungen in SAM 2.0, bspw. die Abschaffung der Solo-Tracking-Durchgänge, verhinderten nicht die Aufdeckung der unterschiedlichen spezifischen Strategien.



Wie bereits in den Voruntersuchungen SAM 1.5 scheint die Rolle des MWBs einen Einfluss auf die Erkennung der beiden unterschiedlichen spezifischen Strategien gehabt zu haben (s. Anhang B.4.3). Diesmal handelt es sich jedoch um die MWB 1, welche diese Manipulation doppelt so oft erkannten. Dieser Unterschied erweist sich als nicht signifikant ( $H_{MWB1} = 12$ ,  $H_{MWB2} = 6$ ,  $p = .180$ ,  $\omega^{69} = .42$ , odds ratio = 2.5,  $1 - \beta = .34$ ). Allerdings kann dieses Ergebnis durch die geringe Teststärke der statistischen Analyse verursacht sein. Für eine sichere statistische Überprüfung ist eine größere Stichprobe nötig. Die dennoch mittlere bis hohe Effektgröße des Unterschieds kann damit erklärt werden, dass VPn häufig versuchen, den Sinn einer Untersuchung zu ergründen (Sarris & Reiß, 2005), was durch ein Orientierungsbedürfnis verursacht wird. Dabei kann aufgrund von Erfahrungen der VPn aus anderen Untersuchungen die Erwartung bestehen, nicht alle oder die richtigen Informationen vom VI erhalten zu haben. Die MWB hatten in deren Untersuchungspausen, während der Probe- und Trainingsphase des anderen MWBs, ausreichend Zeit, um über die tatsächlichen Ziele der Untersuchung zu reflektieren. MWB 1 kannte in seiner Pause bereits die vermeintliche Team-Strategie. Die Tatsache, diese in Abwesenheit des Team-Partners präsentiert bekommen zu haben, könnte zu einem ersten Verdacht geführt haben. Im weiteren Verlauf der Untersuchung könnte die Annahme durch das unterschiedliche Tracking-Verhalten erhärtet worden sein. Die MWB 2 erhielten die spezifische Strategie erst nach deren Pause. Dadurch stand diesen nicht so viel Zeit zum Hinterfragen zur Verfügung.

Abschließend lässt sich feststellen, dass beide MWB den, durch die konträren spezifischen Strategien intendierten, Leistungsunterschied wahrgenommen hatten. Diese subjektive Leistungsdifferenz kann verantwortlich für die vorgefundene negative Veränderung der Stimmung sein. Auf die Selbst- und Fremdbeurteilung der Wichtigkeit ein gutes Tracking-Ergebnis zu erreichen sowie auf die eigene Zufriedenheit mit der Teamleistung der MWB hatte die subjektive Leistungsdifferenz keinen Einfluss. Ein Konflikt, wie das 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) postuliert, scheint bei den MWBn nicht entstanden zu sein. Möglicherweise waren die MWB noch nicht ausreichend motiviert oder der Untersuchungszeitraum war zu kurz, um einen Konflikt aufzubauen. Dafür kann ebenso die hohe Anzahl an MWBn mit verantwortlich sein, welche die Unterschiedlichkeit der spezifischen Strategien erkannten.

---

<sup>69</sup> Konventionen für die Effektgröße  $\omega$  nach Cohen (1988, zitiert nach Sedlmeier & Renkewitz, 2008):  $\omega = .10$  ist ein kleiner,  $\omega = .30$  ist ein mittlerer,  $\omega = .50$  ist ein großer Effekt.

### **3.2.3.4.5 Interpretation und Diskussion einer explorativen Analyse zur Strategie**

Wie eine explorative Analyse zeigte, liegen keine signifikanten Zusammenhänge zwischen der unspezifischen Strategie, erhoben durch den FESE-R (Nachtwei, 2008b), und den objektiven Verhaltensdaten auf Teamebene vor. Diese Ergebnisse bestätigen die Entscheidung, die VPn nicht mehr nach deren unspezifischen Strategie für deren Einsatz in SAM als MWB vorzuselektieren. Die unspezifische Strategie hat beim Solo-Tracking einen Einfluss auf die Objektregelung (Gross & Nachtwei, 2006, 2007), nicht jedoch beim kooperativen Tracking. Scheinbar hat das Versuchssetting die Wirkung der unspezifischen Strategie überlagert. Vorteile einer gemeinschaftlichen Aufgabenbewältigung sind aus anderen Bereichen bekannt, bspw. schneiden Zwei-Personen-Teams in verschiedenen Problem-Solving Situationen besser ab als eine einzelne Personen (Hoc, 2001). Das wird mit sogenannten Synergieeffekten aufgrund von unterschiedlichen Erfahrungen und Kompetenzen begründet (Scherer, 1998). Für die fehlende Wirkung der unspezifischen Strategie auf das kooperative Tracking im Gegensatz zum Solo-Tracking scheint jedoch eine Erkenntnis aus der Joint-Action-Forschung als Erklärungsansatz geeigneter zu sein. Gemeint ist die nicht geplante, unbewusste Synchronisation von Handlungen in sozialen Interaktionen (Knoblich, Butterfill, & Sebanz, 2011). Non-verbales Verhalten scheint diesen Anpassungsvorgang zu erleichtern (Glowinski et al., 2013). Die Leistung eines Teams bei einer gemeinsamen Handlung hängt davon ab, wie gut die einzelnen Personen ihre Handlungen aneinander anpassen können (M. Meyer, Bekkering, Haartsen, Stapel, & Hunnies, 2015). Die beiden MWB konnten mindestens peripher die Joystick- und Körperbewegungen des Teampartners wahrnehmen und sich so an dessen Objektregelungsverhalten anpassen. Eine Möglichkeit diesen Störeinfluss auf die Konfliktbildung zwischen den MWBn zu vermeiden, stellt die räumliche Trennung der MWB dar. Allerdings besteht dabei die Gefahr, dass die MWB die Objektbewegung nicht mehr hauptsächlich der Objektregelung durch den Teampartner und sich selbst zuschreiben (s. Anhang B.4.3), sondern die Ursache für eine schlechte Leistung vermehrt in der Software selbst oder einer Manipulation durch den VI sehen. Dabei könnte es ebenso sein, dass die MWB nicht an die Existenz eines zweiten MWBs glauben, wodurch ein Konfliktaufbau, wie er im 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) beschrieben ist, unmöglich ist.

### **3.2.3.5 Kritik**

Als kritisch bei den Vergleichen der objektiven Tracking-Daten der MWB zwischen SAM 1.0 und 2.0 sowie zwischen dem kurzen und langen Tracking-Pfad sind die unterschiedliche Länge der Tracking-Pfade sowie die Dauer eines Programmschritts der beiden

Softwareversionen anzusehen. Trotz des Versuchs die Daten der aVn auf ein Niveau zu bringen, ist eine Beeinflussung der Ergebnisse nicht auszuschließen.

Ebenso sind die fehlenden demografischen und subjektiven Fragebogendaten als kritisch zu sehen. Von den Untersuchungsbedingungen unabhängige Unterschiede zwischen den beiden Stichproben SAM 1.0 und SAM 2.0 konnten so nicht identifiziert und deren eventueller Einfluss auf die aVn eliminiert werden. Aufgrund der Verwendung der gleichen Ausschlusskriterien<sup>70</sup> für die VPn (s. Abschnitt 3.2.1.2.6.1) konnten einige potentielle Störvariablen ausgeschaltet respektive konstant gehalten werden. Ebenso kann ab einer Stichprobengröße von 20 VPn davon ausgegangen werden, dass eine Äquivalenz der Ausprägungsverteilung der Störvariablen vorliegt (Bortz & Döring, 2006). Dennoch ist eine Beeinflussung der gefundenen Ergebnisse aufgrund von Unterschieden zwischen den beiden Stichproben nicht gänzlich auszuschließen.

### 3.2.3.6 Fazit und Ausblick

In diesem Abschnitt werden die zuvor aufgezeigten Ergebnisse hinsichtlich deren Bedeutung für die Eignung von SAM als Versuchsumgebung zur Aufklärung der weiteren Forschungsfragestellungen des ATEO-Projekts interpretiert.

Die vorgenommenen Veränderungen in SAM 2.0 führten zu einer signifikant ungenaueren und mindestens praktisch bedeutsam schnelleren Objektregelung im Vergleich zu SAM 1.0. Aufgrund der höheren Fehleranzahl in SAM 2.0 besteht ein höherer Bedarf an einem Operateur oder Automaten zur Optimierung des Systemprozesses. Gleichzeitig sind die Operateure und Automaten durch das schnellere Tracking der MWB in ihrer Überwachungs- und Führungstätigkeit mehr gefordert. Dabei kann die gefundene Verringerung der benötigten Tracking-Zeit der MWB von Durchgang zu Durchgang einem möglichen Lerneffekt der Operateure beim Überwachen und Führen von SAM entgegenwirken.

Aufgrund der ermittelten höheren Eignung des langen Tracking-Pfads für die Verwendung in SAM, muss dieser von nun an anstelle des kurzen Pfads in den kooperativen Tracking-Durchgängen von den MWBn bewältigt werden. Dabei wird weiterhin die Position der einzelnen Gabelungen und Hindernisse variiert, um die Komplexität der Pfade aufrechtzuerhalten.

Wie den Ergebnissen zu entnehmen ist, wurden die MWB durch die Rahmenbedingun-

---

<sup>70</sup> Mit Ausnahme der Vorselektion der VPn für die Untersuchung SAM 1.0 nach deren unspezifischen Strategie durch FESE (Nachtwei, 2008a), welche bei den VPn der Untersuchung SAM 2.0 nicht mehr erfolgte. Jedoch konnte mittels der vierten Voruntersuchung „Gabelung ohne Vorselektion der MWB“ (s. Abschnitt 3.2.2) wie auch in dieser Untersuchung nachgewiesen werden, dass die Wirkung der unspezifischen Strategie der VPn auf das kooperative Tracking vernachlässigbar ist.

gen von SAM ausreichend motiviert, die jeweils instruierte spezifische Strategie zum Tracking umzusetzen. Dabei wirkten die beiden konträren spezifischen Strategien weitestgehend, wie im 3K-Modell (Nachtwei, 2008a, s. Abschnitt 2.2.2.1.1) postuliert. Sowohl die Präferenz als auch das Verhalten der MWB beim Tracking wurden in die intendierte Richtung Geschwindigkeit respektive Genauigkeit beeinflusst. Wobei bei dem geschwindigkeitsorientierten MWB 1 noch Optimierungspotential besteht. Ebenso wurde die Leistungsdifferenz von den MWBn wahrgenommen. Allerdings gingen die MWB davon aus, dass es ihrem Teampartner gleich wichtig war, ein gutes Tracking-Ergebnis zu erreichen. Außerdem waren die MWB zum Ende der Untersuchung noch ziemlich zufrieden mit ihrer Teamleistung. Trotz der bereits ausreichenden Motivation der MWB, die jeweilige spezifische Strategie umzusetzen, konnte scheinbar kein Konflikt zwischen ihnen aufgebaut werden. Daher sollte in zukünftigen Untersuchungen mit SAM 2.0 versucht werden, die Motivation der MWB noch mehr zu steigern, bspw. durch eine höhere Differenz zwischen den Stufen der leistungsabhängigen Entlohnung.

Der hohen Anzahl an MWBn, welche die Unterschiedlichkeit der beiden spezifischen Strategien erkannt haben, soll zukünftig durch Ablenkung des jeweils wartenden MWBs während der Trainingsphase des anderen MWBs entgegengewirkt werden. Dafür kann die Erhebung der Kontrollvariablen dienen. Damit lässt sich ebenso der zeitliche Rahmen eines Untersuchungsdurchlaufs kürzen.

Die Verständlichkeit der beiden spezifischen Strategien könnte in zukünftigen Untersuchungen durch kurze Videos von „guten“ (entspricht spezifischer Strategie des jeweiligen MWBs) und „schlechten“ (entspricht spezifischer Strategie des anderen MWBs) Objektregelungen nach der Instruktion der spezifischen Strategie erhöht werden. Dadurch erhalten die MWB Referenzverhalten, mit denen sie ihre eigene Tracking-Leistung vergleichen können. Der VI kann durch Kommentierung der Videos das Augenmerk auf die „schlechte“ Objektregelung legen, so dass den MWBn beim kooperativen Tracking das gegensätzliche Objektregelungsverhalten des Teampartners mehr auffällt und negativ bewertet wird.

Insgesamt sind die gefundenen Ergebnisse vielversprechend, dass SAM 2.0 es ermöglicht, das Verhalten von Operateuren mit dem von Automaten unter kontrollierten Laborbedingungen sinnvoll zu vergleichen. Ob der Optimierungsbedarf des Prozessverlaufs von SAM tatsächlich ausreichend für den Einsatz von Operateuren und Automaten ist, wird sich jedoch erst im Einsatz zeigen können.

Für die Erhebung der subjektiven Beanspruchung der MWB durch die Tracking-Aufgabe kam die  $RSE_A$  (s. Abschnitt 3.2.3.2.2.1) zum Einsatz. Wie sich zeigt, konnte die Skala Unterschiede zwischen den VPn sowie den Pfadlängen aufdecken. Diese kann somit für den

Vergleich von Operateuren und Automaten als Instrument zur Erhebung der subjektiven Beanspruchung der MWB dienen. Die subjektive Beanspruchung der MWB stellt ein Gütemaß der Regelungstätigkeit von Operateuren und Automaten dar.

### **3.3 Von SAM 2.0 zum ATEO Lab System**

Für die Umsetzung des experimentellen Leistungsvergleichs zwischen Entwicklern und Operateuren war es notwendig, SAM um zwei Softwareprodukte zu ergänzen. Zum einen war eine Automatiksoftware (AA-S) nötig, in welcher das Automatikkonzept implementiert werden sollte, welches unter den optimalen Ausprägungen der, im Projekt ATEO gefundenen, Leistungsdeterminanten entstanden ist. Zum anderen benötigten die Operateure einen Arbeitsplatz zur Überwachung und Beeinflussung des Prozesses von SAM, welcher ebenfalls über eine Software umgesetzt wurde. Nachfolgend soll auf beide Softwareprodukte eingegangen werden.

#### **3.3.1 Entwicklung einer Automatiksoftware für SAM**

Um den Leistungsvergleich zwischen Entwicklern und Operateuren vorzubereiten, wurde zum Ende der Entwicklung von SAM 2.0 durch die Autorin mit der Konzeption einer AA-S begonnen. Dafür wurde in einem ersten Schritt eine Hierarchische Aufgabenanalyse (HAA, s. Diaper & Stanton, 2004) durch den Diplomanden der Informatik Kesselring (2009) und die Autorin durchgeführt.

Im Folgenden wird die Entstehung der AA-S dargestellt. Dafür wird kurz auf die HAA eingegangen. Dem folgen die Beschreibungen der AA-S sowie der gebrauchstauglichen Gestaltung des Konfigurationstools der Automatikfunktionen.

##### **3.3.1.1 Hierarchische Aufgabenanalyse für eine Automatiksoftware**

Die HAA diene der Identifizierung der Anforderungen an die AA-S sowie um ein gemeinsames Verständnis von der Software aufzubauen. Nach Marsden und Kirby (2005) ist dieses Verfahren dafür am geeignetsten. Zudem werden bei der HAA kognitive Prozesse außen vor gelassen, so dass für die Anwendung kein fundiertes Wissen über die menschlichen Informationsverarbeitungsprozesse notwendig ist (Hildebrandt et al., 2010). Ebenso können dadurch Aufgaben beschrieben werden, welche von einem Menschen oder einem technischen System, wie einer Automatik, ausgeführt werden (Stanton, 2006). So konnte durch dieses Verfahren eine strukturierte Auflistung der grundlegenden Aufgaben der AA-S in Form von Zielen und Unterzielen, aber auch deren Eigenschaften, wie Priorität, erstellt werden (s. Anhang C.1.1).

Für eine prägnante Beschreibung der Aufgaben wurde die Unterziel-Vorlage zur Prozessüberwachung und -führung sowie eine eigens vorgenommene Erweiterung der Plan-Vorlage von Ormerod und Shepherd (2004) verwendet (Hildebrandt et al., 2010; s. Anhang C.1.1). Als Informationsquellen dienten vorhandene Dokumentationen (wie die von Bothe & Hildebrandt, 2008; Burandt, 2007), SAM 1.5, der Operateursarbeitsplatz, die Assistenzsystemkonzepte aus den Untersuchungen von Krinner (2009) und Interviews von Projektmitgliedern u. a. aus der ersten Projektphase (Kesselring, 2009). Die HAA für die AA-S wurde parallel zu einer HAA für die Operateursarbeitsplatzsoftware (OA-S, s. Abschnitt 3.3.2) durchgeführt, wobei die letzten Iterationsschleifen in enger Verknüpfung miteinander zum Abgleich durchgeführt wurden.

Mittels der obersten Hierarchieebenen konnten Use Cases erstellt werden. Informatiker arbeiten für die Analyse der Anforderungen an ein zu entwickelndes System mit Use Cases, vor allem bei der objektorientierten Softwareentwicklung (Balzert, 2000; Sommerville, 2007). Use-Case-Diagramme bieten für die Implementierung durch das höhere Abstraktionsniveau (Balzert, 2000) im Vergleich zu den Aufgabemodellen der HAA eine gute Übersicht auf das zukünftige System. Zudem diene das Ergebnis der HAA bei der Konzeption und Implementierung der Architektur der AA-S inklusive der ersten Automatikfunktionen als eine Art Checkliste zur Prüfung der Vollständigkeit (Hildebrandt et al., 2010).

Eine Beschreibung der HAA, deren Anwendung und Ergebnis im Projekt ATEO findet sich in Kurzform im Anhang C.1.1 und ausführlich in Hildebrandt et al. (2010).

### **3.3.1.2 Konzeption und Implementierung einer Automatiksoftware**

Die Softwareimplementierung wurde in den, von der Autorin betreuten, Diplomarbeiten von Kesselring (2009) begonnen und von Hasselmann (n. D.) zum ATEO Automation Framework (AAF) weitergeführt. Das AAF stellt den technischen Rahmen dar, über den Automaten Einfluss auf SAM haben. Dafür wird dem AAF bei jedem SAM-Programmschritt (ca. alle 39 ms) der aktuelle Zustand von SAM übermittelt (Kosjar, 2011). Dieser enthält Informationen wie die aktuellen Joystick-Eingaben der MWB und die Abstände des Tracking-Objekts in Pixeln zu einzelnen Pfadelementen. Dieser SAM-Zustand kann von jeder einzelnen Automatikfunktion manipuliert und um weitere Daten ergänzt werden, wie eine anzuzeigende Bilddatei an einer bestimmten Position auf dem Tracking-Pfad im Sichtbereich der MWB. Der veränderte SAM-Zustand wird wieder zurück an SAM übermittelt und dort umgesetzt.

Neben dem Leistungsvergleich soll die AA-S gemeinsam mit SAM über das Projekt ATEO hinaus anderen Forschern zur Aufklärung unterschiedlichster Fragestellungen dienen. Daher wurde darauf geachtet, dass diese flexibel, einfach und schnell an die jeweils

bestehenden Anforderungen angepasst und um beliebige Funktionen erweitert werden kann. Ebenso wurde eine saubere Trennung zu SAM sowie eine Vernetzungsmöglichkeit umgesetzt, um so wenig wie möglich Änderungen im Quellcode von SAM vornehmen zu müssen (Niestroj, 2009).

Die ersten Automatikfunktionen leitete die Autorin aus den Ergebnissen der HAA für die AA-S (s. Anhang C.1.1) und die OA-S (s. Anhang C.1.2) als auch von den experimentell gewonnenen Assistenzsystemkonzepten in der ersten ATEO-Projektphase (s. Krinner, 2009) ab. Hierbei handelte es sich um Funktionen, welche den MWBn visuelle oder auditive Hinweise auf herannahende Pfadelemente oder Empfehlungen hinsichtlich des Objektregelungsverhaltens vermitteln. Eine dieser Funktionen weist auf ein vorausliegendes Hindernis hin, durch Einblendung eines Hindernis-Icons am Pfadrand der betroffenen Pfadseite, kurz bevor das Hindernis im Sichtbereich der MWB erscheint. Ebenso wurden bereits Funktionen mit direktem Zugriff auf den MWB-Objekt-Tracking-Umgebung-Regelkreis vorgesehen, bspw. um die Objektgeschwindigkeit in Kurven zu verringern, wenn die Gefahr bestand, den Pfad zu verlassen. Nachdem erfahrene Entwickler-VPn 30 Automatikkonzepte in der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“ der Autorin erstellt hatten (s. Abschnitt 5.2.3), sollte von diesen für den Leistungsvergleich mit einem Operateur das Konzept mit der höchsten Güte in die AA-S integriert werden. Dadurch können die Automaten dieses Konzepts den Prozess von SAM überwachen und führen. Dies sollte der Aufklärung einer der Forschungsfragestellungen des ATEO-Projekts dienen, wie gut Entwickler den Bedarf an Automaten sowie deren notwendigen Funktionalitäten antizipieren können unter den optimalen Ausprägungen ihrer Leistungsdeterminanten. Ebenso sollte durch den Leistungsvergleich mit einem Operateur bei der Überwachung und Führung des SAM-Prozesses die Fragestellung aufgeklärt werden, welche Funktionen besser durch einen Operateur und welche besser durch eine Automatik ausgeübt werden sollten.

Während der Zeit der Gütebewertung der 30 Automatikkonzepte sollte bereits eine maximale Anzahl der darin enthaltenen nützlichen Funktionen in die AA-S für die Nachnutzbarkeit durch andere Forscher implementiert werden. Dafür wurden die Funktionen mittels eines, von der Autorin erstellten, Fragebogens (s. Anhang C.2) von fünf Diplomanden der Informatik dahingehend bewertet, wie leicht, aufwändig und nützlich deren Implementierung in die AA-S ist. Der Fragebogen wurde in Anlehnung an Parasuraman et al. (2000) entwickelt, nach denen zur Bewertung von Automationskonzepten die Leichtigkeit und die Kosten von deren Implementierung als Kriterien herangezogen werden können. Durch dieses Vorgehen konnten geeignete Funktionen identifiziert werden (Wickert, 2013). Außerdem zeigte sich, dass die

meisten Automatikkonzepte von den Entwickler-VPn für eine Implementierung unzureichend spezifiziert sind. Daher musste die Autorin die darin enthaltenen Funktionen konkretisieren und weiterentwickeln, ohne deren Grundideen zu verändern. Zudem ähneln sich viele Funktionen, wie die Empfehlung einer Gabelungsalternative per Pfeil, Einfärbung des relevanten Pfads oder auditiver Richtungsempfehlung. Daher wurden ähnliche Funktionen als eine Funktion mit Parametern zum Konfigurieren verschiedenster Eigenschaftsausprägungen umgesetzt (Kosjar, 2012). Dadurch ist es möglich, geeignete Ausprägungen der Parameter sowohl beim Tracking als auch empirisch zu ermitteln.

Die Autorin hat bis 2012 die Implementierung einer Vielzahl dieser Funktionen durch Informatik-Diplomanden angeleitet und deren Funktionalität manuell getestet. Dabei wurden die Funktionen zur visuellen und auditiven Vermittlung von Ist- und Ideal-Werten an die MWB durch Seid (2012a) implementiert, wie die Anzeige der Ideal-Tracking-Objektgeschwindigkeit. Visuelle und auditive Hinweise, Warnungen vor Pfadelementen und Empfehlungen für den Umgang mit diesen wurden durch Kosjar (2012) umgesetzt, wie ein Warnton bei Erreichen eines kritischen Abstands des Tracking-Objekts zum Pfadrand. Funktionen, welche direkt den MWB-Objekt-Tracking-Umgebung-Regelkreis beeinflussen, wurden von Wickert (2013) umgesetzt, wie das automatische Regeln des Tracking-Objekts zurück auf den Tracking-Pfad. Funktionen zur Berechnung der Ideal-Werte, wie für die Joystick-Eingaben oder die Tracking-Trajektorie, wurden durch Weidner-Kim (2014, 2015) implementiert. Damit die Aktivitäten der Automatikfunktionen während eines Tracking-Durchgangs erfasst sowie im Nachgang ausgewertet werden können, wurde durch Seid (2012b) das Logfile von SAM und das LFA-Tool zur Version 2.0 erweitert.

### **3.3.1.3 Graphical User Interface der Automatiksoftware**

Für den Einsatz der AA-S in experimentellen Untersuchungen müssen die Automaten durch einen VI erstellt und konfiguriert werden können, der zumeist keine Programmierkenntnisse aufweist. Deshalb wurde für das AAF ein grafisches Konfigurationstool der Automatenfunktionen durch die Autorin konzipiert, welches nachfolgend vereinfacht als Graphical User Interface (GUI) der AA-S bezeichnet wird. Dieses ermöglicht es, die verschiedensten Funktionen des AAFs zur Überwachung und Führung von SAM einfach und schnell zu konfigurieren und zu Automaten zusammenzustellen. Das AAF und das GUI bilden die AA-S. Diese wurden in Squeak programmiert. In seinen Grundzügen wurde das GUI durch die Diplomandin Fuhrmann (2010) implementiert und in der Studien- und Diplomarbeit von Kosjar (2011, 2012) ausgebaut.



Das GUI-Layout teilt sich in vier Hauptbereiche auf (s. Abbildung 18), welche nachfolgend kurz vorgestellt werden (Details s. Fuhrmann, 2010; Kosjar, 2011, 2012; Wickert, 2013).

### 1. Menüleiste

Horizontal im oberen Bereich des GUIs befindet sich die Menüleiste. In dieser werden dem Benutzer aus anderen Programmen gewohnte Funktionalitäten angeboten. So kann mittels der vier Buttons auf der linken Seite eine Automatik neu erstellt, geöffnet, gespeichert oder unter neuem Namen gespeichert werden. Auf der rechten Seite befindet sich der Button zum Beenden des Programms. Dieser wurde maximal weit entfernt von den anderen Buttons positioniert, um die Gefahr einer versehentlichen Bedienung zu minimieren.

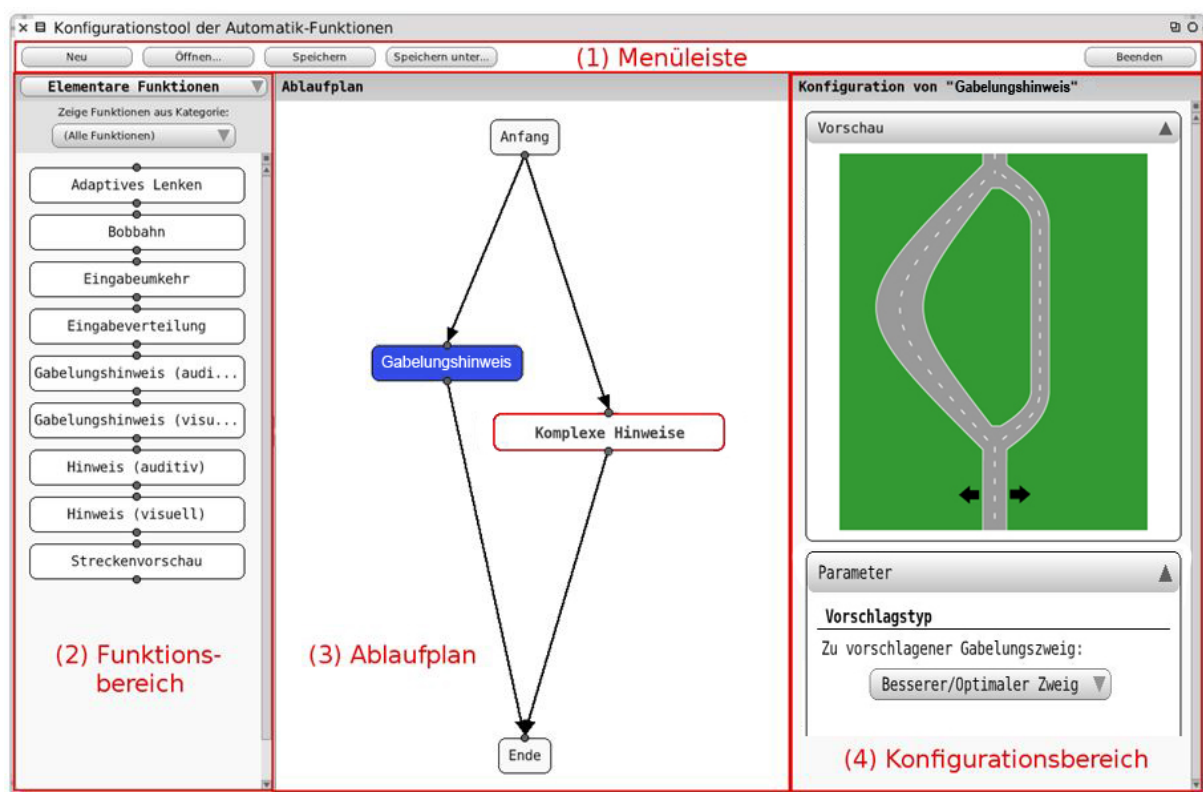


Abbildung 18: Graphical User Interface der ATEO Automatiksoftware

### 2. Funktionsbereich

Links befindet sich der Funktionsbereich. In diesem stehen dem Benutzer aus Gründen der Übersichtlichkeit zwei Ansichten zur Verfügung. Entweder kann eine Liste von elementaren Funktionen oder aber eine Liste von bereits kombinierten Funktionen über ein Dropdown-Menü zur Anzeige ausgewählt werden. Bei den kombinierten Funktionen handelt es sich um Funktionen, welche unter einem gemeinsamen Namen als Automatik abgespeichert sind. Die

einfachen und kombinierten Funktionen sind alphabetisch angeordnet. Zudem können diese über ein Dropdown-Menü nach Kategorien gefiltert werden.

### **3. Ablaufplanbereich**

In der Mitte liegt der Ablaufplanbereich. Hier werden die einzelnen Funktionen miteinander in Form eines Ablaufplans mit einem Anfang und einem Ende zu einer Automatik kombiniert.

### **4. Konfigurationsbereich**

Rechts befindet sich der Konfigurationsbereich der Funktionen. In diesem Bereich kann einer neu erstellten Automatik ein frei wählbarer Name, eine kurze Beschreibung sowie Kategorien für den Funktionsbereich gegeben oder diese angepasst werden.

Wird eine einfache Funktion im Ablaufplanbereich per Links-Mausklick ausgewählt, erscheinen deren Einstellmöglichkeiten in diesem Konfigurationsbereich. Es kann für jede Funktion einer Automatik deren Aktivierungsbedingungen definiert werden, d. h. ob diese über den gesamten Tracking-Pfad oder nur bei bestimmten Ereignissen aktiv sein soll. Unter Ereignisse fallen das Auftreten einzelner Pfadelemente, wie eine Gabelung, aber auch bestimmter Objektregelungsverhalten eines oder beider MWB, wie zu starke Joystick-Auslenkung. Für jedes auf ein Pfadelement bezogenes Ereignis kann dessen Anfang und Ende durch Abstände in Pixeln zum Beginn des Pfadelements definiert werden. Solange sich das Tracking-Objekt in diesem Aktivitätsbereich befindet, ist die Funktion aktiv. Darüber hinaus können im Konfigurationsbereich weitere Eigenschaften der Funktionen eingestellt werden. Für Funktionen zur visuellen Informationsvermittlung an die MWB kann eine Bilddatei im PNG-Format ausgewählt werden. Für jede Bilddatei kann u. a. eine Blinkfrequenz durch Definition der Einblend- und Ausblenddauer in ms eingestellt werden. Im Konfigurationsbereich existiert über dem Parameterbereich ein Vorschaubereich. Hier werden die ausgewählte Bilddatei und die Auswirkungen der weiteren Parametereinstellungen dargestellt. Für Funktionen zur auditiven Informationsvermittlung an die MWB kann die abzuspielende MP3-Datei ausgewählt und zur Überprüfung der richtigen Auswahl abgespielt werden. Für eine direkt in den MWB-Objekt-Tracking-Umgebung-Regelkreis eingreifende Funktion kann u. a. der prozentuale Wert der Objektregelungsübernahme festgelegt werden. Viele Parameter hängen von der konkreten Funktionalität einer Funktion selbst ab.

Hat der Benutzer im Ablaufplan eine kombinierte Funktion per Links-Mausklick markiert, steht ihm im Konfigurationsbereich die Funktionalität „Detailgrad erhöhen“ zur

Verfügung. Dadurch werden deren einzelne Funktionen als Ablaufplan angezeigt. Diese einzelnen Funktionen können wiederum durch Auswahl per Links-Mausklick konfiguriert werden. Möchte der Benutzer zur ursprünglichen Anzeige zurückwechseln, steht diesem nun im Konfigurationsbereich die Funktionalität „Detailgrad verringern“ zur Verfügung.

#### **3.3.1.3.1 Gebrauchstaugliche Gestaltung des GUIs**

Bei der GUI-Gestaltung wurden software-ergonomische Prinzipien, Empfehlungen, Heuristiken, Richtlinien und Normen berücksichtigt. Daher wurde ein minimalistisches (s. Nielsen, 1994) und strukturiertes (s. Preim, 1999) Design gewählt, um eine Übersichtlichkeit und insbesondere einfache Orientierung für die Benutzer zu erreichen (s. Herczeg, 2009). Dafür wurden zusammengehörige Elemente nach dem Gesetz der Nähe gruppiert (s. Wertheimer & Köhler, 1952, zitiert nach Sarodnick & Brau, 2006; DIN EN 894-1, 1997), so dass sich die vier Hauptbereiche, aber auch innerhalb des Konfigurationsbereichs einzelne Parameterbereiche ergaben.

Um eine hohe Aufgabenangemessenheit des GUIs zu erreichen, enthält dieses nur für die Erstellung und Konfiguration von Automaten relevante Elemente (s. DIN EN ISO 9241-110, 2006). Ebenso existieren aufgabenangemessene Voreinstellungen. So liegen bspw. für Funktionen, welche nur für bestimmte Pfadelemente bestimmt sind, u. a. zum Hinweisen auf ein Hindernis, bereits vorkonfigurierte Aktivitätsbereiche vor.

Um Missverständnisse bei der Bedienung des GUIs zu vermeiden, wurde eine leichte und unmittelbar verständliche Gestaltung dieser angestrebt (s. DIN EN 894-1, 1997; Johannsen, 1993). Dafür wurden einfach erkennbare und verständliche Begriffe als Bezeichner verwendet (s. DIN EN ISO 9241-110, 2006; DIN ISO/IEC 12119, 1995), welche das zugrunde liegende Referenzobjekt gut und unverwechselbar beschreiben (s. Wandmacher, 1993). Ebenso sind Bereichs- und Funktionsbezeichner, die verschiedenen Bedienelemente sowie deren Status unterschiedlich gestaltet, damit diese differenzierbar sind (s. Prinzip der Unterscheidbarkeit der DIN EN 894-1, 1997; Herczeg, 2009; Preim, 1999). So werden markierte Funktionen visuell hervorgehoben durch eine weiße Schrift des Bezeichners auf blauem Hintergrund, wohingegen die Bezeichner nicht markierter Funktionen eine schwarze Schrift auf weißem Hintergrund aufweisen. Darüber hinaus wurde das GUI selbstbeschreibend umgesetzt (s. DIN EN ISO 9241-110, 2006), indem der Benutzer über Tooltips Beschreibungen einzelner Bereiche und Elemente erhält, sobald der Mauscursor über diese wandert (s. aktive Hilfe bei Herczeg, 2009).

Durch die Verwendung von einheitlichen Bedienelementen und -möglichkeiten sollte eine hohe Erwartungskonformität umgesetzt werden (s. DIN EN 894-1, 1997; DIN EN ISO 9241-110, 2006; Nielsen, 1994; Sarodnick & Brau, 2006). So können bspw. die Positionen der

einzelnen Funktionen im Funktions- und Ablaufplanbereich, der Ereignisse im Aktivierungsbereich oder der Bilder im Vorschaubereich per „Drag & Drop“ mittels der linken Maustaste verändert werden (s. generische Funktionen und innere Konsistenz bei Herczeg, 2009). Ebenso wurde die Erfahrung der Zielgruppe berücksichtigt (s. Wickens, Lee, Liu, & Gordon-Becker, 2004), indem das Layout und die Bedienmöglichkeiten an für Forscher gängige Programme orientiert wurden, wie Adobe oder Microsoft Office (s. äußere Konsistenz bei Herczeg, 2009).

Eine lernförderliche Gestaltung des GUIs soll u. a. durch die Beschreibungen der Handlungsmöglichkeiten geschaffen werden (s. DIN EN ISO 9241-110, 2006). Der Benutzer wird bspw. im Konfigurationsbereich darauf hingewiesen kombinierten Funktionen einen Namen, eine Beschreibung und Kategorien für den Funktionsbereich geben zu können.

Das GUI soll eine hohe Steuerbarkeit dadurch aufweisen, dass der Benutzer jederzeit die Erstellung und Konfiguration einer Automatik starten, unterbrechen sowie die Reihenfolge der Bedienschritte vorrangig selbst bestimmen kann (s. DIN EN ISO 9241-110, 2006). Es wurde ein benutzergesteuerter Dialog umgesetzt (s. Herczeg, 2009). So ist es unerheblich welche Funktion sowie welcher Parameter zuerst bearbeitet wird. Weiterhin sind im Konfigurationsbereich Parameter inhaltlich gruppiert in Bereiche, welche Default minimiert sind. Bei Bedarf kann sich der Benutzer diese in die Anzeige per Links-Mausklick auf das Öffnungssymbol<sup>71</sup> rechts neben dem jeweiligen Bereichsbezeichner holen.

Zudem sind im GUI Warnungen bei fehlerhaften Eingaben vorgesehen, wodurch eine hohe Fehlertoleranz erreicht werden soll (s. DIN EN ISO 9241-110, 2006). Werden bspw. nicht umsetzbare oder unvollständige Einstellungen bei den Aktivierungsereignissen einer Funktion vorgenommen, erhält der Benutzer in der Kopfzeile des Aktivierungseinstellungsbereiches die Warnung „Bedingung ungültig“.

Um die Fähigkeiten, Erfahrungen und Vorlieben der unterschiedlichen Benutzer abzudecken, wurden Individualisierbarkeiten im GUI vorgesehen (s. DIN EN ISO 9241-110, 2006). So bietet dieses alternative Bedienwege für eine Aktion an, bspw. kann der Detailgrad einer kombinierten Funktion im Ablaufplan im Konfigurationsbereich per Button, per Doppelklick mit der linken Maustaste auf den Funktionsbezeichner im Ablaufplan oder im Kontextmenü per rechter Maustaste auf den Funktionsbezeichner im Ablaufplan verändert werden. Ebenso kann die Breite der Hauptbereiche durch „Splitter“ zwischen Funktionsbereich und Ablaufplanbereich sowie Ablaufplanbereich und Konfigurationsbereich verändert werden.

Weitere gewählte Gestaltungsmerkmale zur Erreichung eines gebrauchstauglichen

---

<sup>71</sup> Dreieck-Symbol mit einer Spitze senkrecht nach unten bedeutet ein Bereich kann geöffnet werden.

GUIs sind, aus Platzgründen, dem Anhang C.3 zu entnehmen.

Das GUI wurde von Kosjar (2012) einer empirischen Usability Evaluation unterzogen. Dabei konnten keine gravierenden Mängel aufgedeckt werden. Viele der gefundenen Probleme bzw. Wünsche wurden behoben, bspw. ließ sich ein Dropdown-Menü nicht öffnen. Ebenso wurden gewünschte Bedienmöglichkeiten umgesetzt, wie die Erstellung einer Verbindung zwischen zwei Funktionen im Ablaufplan durch Links-Mausklick auf den Verbindungspunkt der einen Funktion und ziehen auf den Verbindungspunkt der anderen Funktion mit anschließendem Loslassen der Maustaste. Der Lösungsaufwand einiger Wünsche überstieg den Nutzen und wurde daher nicht umgesetzt, wie die Auswahlmöglichkeit von unterschiedlichen Schrifttypen.

### **3.3.2 Entwicklung eines Operateursarbeitsplatzes für SAM**

Auf die Entwicklung eines adäquaten Operateursarbeitsplatzes (OAs) für die Prozessüberwachung und -führung von SAM wird im Folgenden nur kurz eingegangen, weil dieser im Rahmen dieser Arbeit eine untergeordnete Rolle spielt (Details s. Anhang D; Bothe et al., 2009; Burandt, 2007; Hildebrandt, 2008; Leonhard, 2010, 2013; Nachtwei, 2011a, 2011b; Nachtwei & Meyer, 2011; Niestroj, 2009; H. Schwarz, 2009; Stade, 2012).

In der ersten ATEO-Projektphase wurde bereits ein sehr einfach gehaltener OA entwickelt. Dieser bestand aus zwei Monitoren (Monitor 1: Echtzeitvideo von MWBn, Monitor 2: Tracking-Umgebung) und einer Tastatur zur visuellen und, über Kopfhörer, auditiven Informationsvermittlung an die MWB (Burandt, 2007). Dieser OA wurde in der zweiten Projektphase von Nachtwei (2011a) weiterentwickelt. Dafür wurde in einem Team, zu dem die Autorin gehörte, ebenfalls eine HAA (Diaper & Stanton, 2004) iterativ hinsichtlich der Aufgaben eines Operateurs bei der Überwachung und Führung von SAM durchgeführt (Details s. Hildebrandt et al., 2010). Aufbauend auf den Ergebnissen dieser HAA (s. Anhang C.1.2), Erkenntnissen aus der Literatur, Interviews realer Operateure sowie mittels heuristischer Evaluationen (s. Nielsen & Molich, 1990) mit Usability Experten entstand iterativ ein erster Prototyp mit einem GUI für einen 30“ Monitor (Nachtwei, 2011a). Dieser bietet dem Operateur mehr Einblicke in den Prozess von SAM als der OA der ersten Projektphase, wie bspw. eine Pfadvorschau. Durch sukzessive Erhöhung der Funktionalitäten entstanden zwei weitere Varianten, wobei die dritte Version die höchste Anzahl an Informationen über den SAM-Prozess und Beeinflussungsmöglichkeiten von diesem anbietet (Details s. Anhang D; Nachtwei & Meyer, 2011). Mittels dieser beiden Varianten kann der Operateur erstmalig Eingriffe zur direkten Beeinflussung des Tracking-Objekts oder des Joystick-Inputs der MWB ausführen (Nachtwei, 2011b). Die drei OA-Varianten wurden hinsichtlich deren Eignung für den experimentellen Leistungsvergleich

von Entwicklern und Operateuren experimentell evaluiert (Nachtwei, 2011a). Am geeignetsten erwies sich die dritte Variante. Bei dieser empfanden die MWB die Tracking-Aufgabe signifikant am wenigsten anstrengend, verringerte sich die mentale Beanspruchung und erhöhte sich das Situationsbewusstsein des Operators (Nachtwei & Meyer, 2011). Durch eine Replikationsstudie konnten diese Befunde untermauert werden (Stade, 2012). Bei dem so entwickelten OA handelt es sich um eine Art Leitwarte, welcher im Projekt als ATEO Master Display (AMD) bezeichnet wird (Nachtwei, 2011a).

### **3.3.2.1 Entwicklung des Operateursarbeitsplatzes für 13,3“ Monitor**

Im Rahmen der, durch die Autorin durchgeführten, Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“ mit professionellen Entwicklern sollte der OA diesen als eine Informationsquelle dienen (s. Abschnitt 3.3.2). Aufgrund der mit der Untersuchung verbundenen Reisen sowie der Verwendung des einzigen 30“ Monitors in den zeitgleich stattfindenden Evaluations-Untersuchungen der OA-Varianten (s. Abschnitt 3.3.1.2) musste die OA-Software (OA-S) auf einem Laptop mit einem 13,3“ Monitor und einer Auflösung von 1280x800 Pixeln angewendet werden können. Im Zuge der Entwicklung der OA-S ging, trotz intensiver Bemühungen mittels des Versionsverwaltungssystems Git und regelmäßigen Besprechungen mit allen Beteiligten, die Anforderung der beliebigen Skalierbarkeit an das GUI der OA-S verloren. So stellte sich kurze Zeit vor dem Start der Untersuchung mit den Entwicklern heraus, dass die Elemente des GUIs mit pixelgenauen Positionen und fixen Größen für einen 30“ Monitor mit einer fixen Auflösung von 2560x1600 Pixeln implementiert worden waren.

Die Anforderung der beliebigen Skalierbarkeit an das GUI der OA-S bestand auch aufgrund der Nachnutzbarkeit der OA-S durch andere Forschergruppen. Diese Anforderung konnte im Rahmen der noch verbliebenen Zeit nicht umgesetzt werden. Hauptsächlich, weil für einzelne Komponenten keine geeigneten Lösungen gefunden werden konnten, etwa der Pfadvorschau. Deshalb wurde von Hasselmann (2010), unter Betreuung durch die Autorin, die OA-S so angepasst, dass weiterhin nur eine Version jedoch mit Wechselmöglichkeit zwischen zwei Größenvarianten des GUIs existiert: 30“ und 13,3“. Dies gewährleistet, dass künftige Veränderungen bei einer der beiden GUI-Varianten in den meisten Fällen automatisch zu Veränderungen bei der anderen GUI-Variante führen. Die Funktionalität der OA-S ist für beide Layouts gleich gehalten (s. Anhang D.4). (Details s. Hasselmann, 2010)

### 3.3.3 ATEO Lab System (ALS)

Die in der zweiten ATEO-Projektphase entwickelten Software-Komponenten SAM 2.0, OA-S, AA-S und LFA 2.0 bilden das ATEO Lab System (ALS), welches dem experimentellen Leistungsvergleich zwischen Entwicklern, über deren Automaten, und Operateuren dient. Die Programmierumgebung Squeak stellt hohe Anforderungen an die CPU-Rechenleistung eines Computers, um die Morphe in einer angemessenen Geschwindigkeit bewegen zu können. Allerdings wird dabei nur ein CPU-Kern verwendet. Neben SAM benötigt die OA-S eine hohe CPU-Rechenleistung. Damit unter Einsatz der OA-S die Programmschritte von SAM sowie die Aufzeichnung aller relevanten Daten in ein Logfile weiterhin in einer hohen Taktfrequenz (ca. alle 39 ms) möglich sind, besteht das ALS aus einem Computernetzwerk.

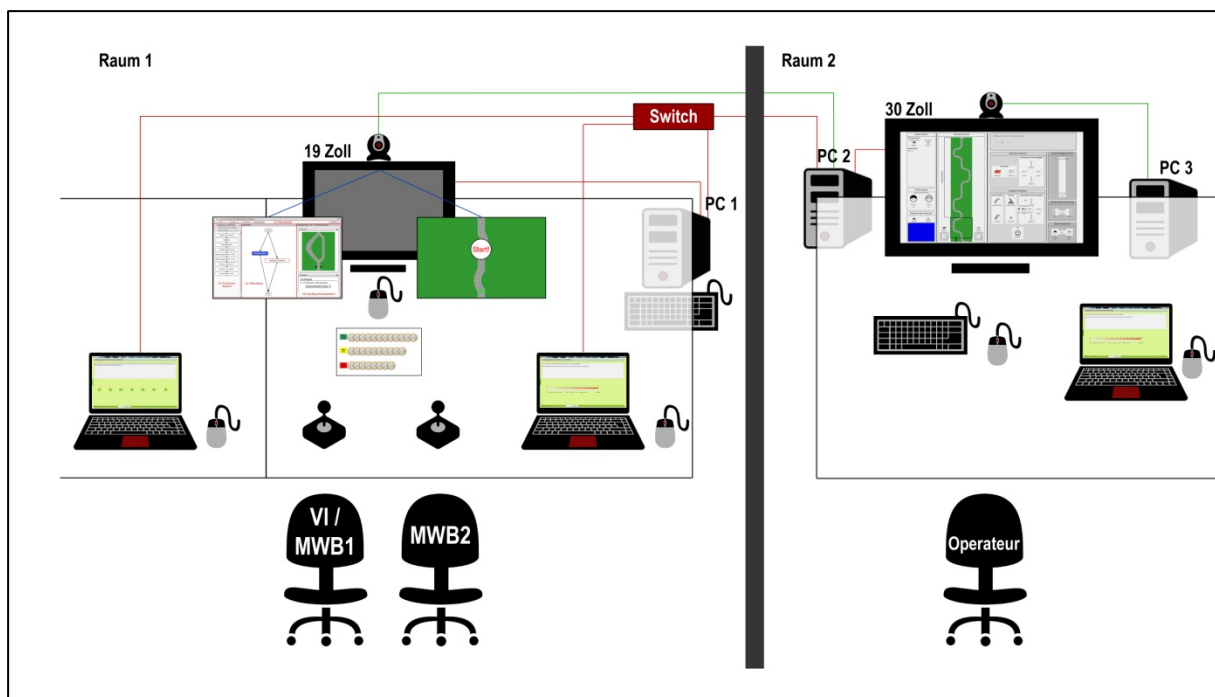


Abbildung 19: Schematische Zeichnung des ATEO LAB Systems (ALS)

Über einen Ethernet-Switch (5-Port10/100M) sind ein Computer für SAM und die AA-S und ein zweiter Computer für die OA-S und die Videoaufnahmen der MWB miteinander verbunden (s. Abbildung 19). Auf einem dritten Computer werden die Videoaufnahmen vom Operateur gespeichert. Ein eigener Computer für die OA-S wurde auch gewählt, damit die MWB der Vergleichsuntersuchung „Operateur vs. Automaten“ durch einen räumlichen Abstand nicht wahrnehmen können durch wen oder was sie unterstützt werden. Alle MWB sollen in der Vergleichsuntersuchung die Information erhalten, durch Automaten unterstützt zu werden. Dieses Vorgehen war geplant, weil die Unterstützung von einem Menschen vermutlich

anders akzeptiert und berücksichtigt wird als von einem technischen System. Das allein könnte die Leistung der MWB beeinflussen.

Im ALS können die OA-S und die AA-S gemeinsam oder einzeln zur Beeinflussung von SAM 2.0 eingesetzt werden. Zusätzlich können für Befragungen der MWB oder des Operators zwischen den Durchgängen und am Schluss Laptops eingesetzt werden. Insgesamt besteht das ALS somit aus einem Aufbau von bis zu sechs Computern.

### 3.4 Fazit zur Umsetzung der Komponenten des ATEO Lab Systems in Squeak

Die Autorin leitete drei Jahre während ihrer Tätigkeit bei prometei und weitere ein- und halb Jahre neben einer neuen beruflichen Tätigkeit intensiv die Implementierung von SAM 1.5 und 2.0, dem SAM- und dem ALS-Logfile, dem LFA-Tool 1.0 und 2.0, der AA-S (AAF und GUI) sowie der mobilen Variante der OA-S im Rahmen von 19 Studien- und Diplomarbeiten an. Diese Aktivität beinhaltete u. a. die Konzeption, regelmäßige Besprechungen mit den Programmierern sowie die manuelle Testung der einzelnen Komponenten.

Dass die Programmierumgebung Squeak prinzipiell für die Umsetzung von SAM geeignet ist, konnte Gross (2008) nachweisen. Allerdings kamen im Verlauf der zweiten ATEO-Projektphase mit ansteigender Komplexität der Mikrowelt und den weiteren Softwareprodukten einige Einschränkungen von Squeak zum Tragen, wie die benötigte hohe CPU-Rechenleistung. Zudem wird Squeak aktuell nur von einer kleinen Community beherrscht, wodurch nur wenige Ansprechpartner mit Expertise und für auftretende Probleme nicht immer Lösungen existieren. Zugleich wies die Architektur von SAM durch die umfangreichen Veränderungen und Neuerungen einige Defizite auf, so existierten viele zyklische Abhängigkeiten und es wurden globale Variablen verwendet (s. Hildebrandt, 2008, 2009; Niestroj, 2009). Dadurch waren die Wiederverwendbarkeit, die Wartung und die Erweiterbarkeit von SAM mit hohen zeitlichen Kosten verbunden. Aufgrund des inhaltlichen und zeitlichen Fortschritts sowie der personellen Ressourcen des ATEO-Projekts in der zweiten Phase konnte eine Neuimplementierung<sup>72</sup> zur Aufklärung der Fragestellungen nicht schnell genug umgesetzt werden. Daher wurden zum damaligen Zeitpunkt nur einzelne chaotische Bereiche und der umfangreiche tote

---

<sup>72</sup> Trotzdem wurde für die Nutzbarkeit durch andere Forschergruppen unter Betreuung der Autorin eine erste rudimentäre SAM-Variante in Java implementiert, welche die zuvor angesprochenen Mängel von SAM<sub>Squeak</sub> nicht aufweist und für ein besseres Verständnis eine hierarchische Schichtenarchitektur besitzt (s. Hildebrandt, 2008) und bei einem Performance-Vergleich SAM<sub>Java</sub> mit SAM<sub>Squeak</sub> besser abschneidet (s. Hildebrandt, 2009). Diese SAM<sub>Java</sub>-Variante wurde zu einem "Generic Laboratory System" ausgebaut (Wagner, 2012a, 2012b). Diese Software-Produktlinie stellt eine Plattform für einen Softwaretypen dar, dessen Funktionalitäten teilweise variabel sind. Indem für diese variablen Stellen bereits verschiedene Komponenten existieren, kann eine bestimmte Simulation schnell zusammengesetzt sowie angepasst werden und nur einzelne Komponenten müssen ggf. neu implementiert werden (Washington, 2013). Neben SAM kann bereits eine Lotsen-Schlepper-Simulation umgesetzt werden.



Code beseitigt sowie die Verwendung von globalen Variablen verringert (s. Niestroj, 2009).

Nichtsdestotrotz hatte das in Squeak programmierte ATEO Lab System bis zum Ende der zweiten ATEO-Projektphase eine ausreichende Komplexität und Nachnutzbarkeit zur Aufklärung von weiteren Fragestellungen aus dem Human Factors Bereich erreicht. Die implementierten Automatikfunktionen können nicht nur der Aufklärung der Fragestellungen des ATEO-Projekts hinsichtlich der Leistung der verschiedensten Automationsvarianten bei der Prozessüberwachung und -führung von SAM und deren Wirkung auf die MWB (bspw. mentale Beanspruchung, Situationsbewusstsein, Vertrauen) dienen. Ebenso kann bspw. die Wirkung verschiedenster Informationsvermittlungsarten (wie synchron/asynchron, räumlich visuell/auditiv) getestet werden. In der dritten ATEO-Projektphase wurden die Einsatzmöglichkeiten durch die Implementierung von kooperativen Automationen zudem erweitert (von Bernstorff, 2014).

Ebenso kann das ALS oder dessen einzelne Softwarekomponenten in Untersuchungen anderer Disziplinen nützlich sein, wie der Persönlichkeitsdiagnostik, Sozialpsychologie, Biologischen Psychologie, Allgemeinen Psychologie, Kognitiven Psychologie (Niestroj, 2009). Der zukünftige Einsatz über das ATEO-Projekt hinaus wird durch die umfangreichen schriftlichen Dokumentationen und Quellcodekommentare der Softwarekomponenten, deren GUIs zur Konfiguration sowie Möglichkeiten zur Erweiterung unterstützt. Die Dokumente und ALS Softwarekomponenten können bei der Autorin angefragt werden.

## **PART II**

### **Determinanten von Entwicklern**

## 4. Theoretischer Hintergrund

### 4.1 Automation in Mensch-Maschine-Systemen

Automatisierung ist längst nicht mehr Zukunft, sondern allgegenwärtig. Bei immer mehr Aufgaben wird der Mensch<sup>73</sup> in fast allen Lebensbereichen durch automatisierte (Teil-) Systeme unterstützt oder sogar ersetzt. Es existiert eine nahezu unendliche Anzahl an Anwendungsgebieten (Sarter et al., 1997). Aufgaben, welche heute noch allein vom Menschen ausgeführt werden, können zukünftig automatisiert sein (Hauß & Timpe, 2000).

In diesem Abschnitt sollen die Begrifflichkeiten „Automatisierung“, „Automation“, „Automatik“, „Assistenzsystem“ und „Funktion“ erläutert werden. In der Literatur lassen sich dazu unterschiedliche Definitionen finden. Nicht immer wird dabei eine klare Abgrenzung vorgenommen.

#### 4.1.1 Begriffsbestimmung „Automatisierung“, „Automation“ und „Funktion“

Die Begriffe „Automatisierung“ und „Automation“ sind abgeleitet von dem „griechischen Wort Automatos, was so viel wie sich selbst bewegend, aus eigenem Antrieb, bedeutet“ (Hauß & Timpe, 2000, S. 42). Hauß und Timpe (2000) behandeln beide Begriffe synonym, aufgrund ihrer historischen Verwendung in Forschung und Praxis. Sie umfassen sowohl den Vorgang der Übergabe einer Aufgabe vom Menschen an ein technisches System als auch das Ergebnis dieses Prozesses selbst. Nachstehend wird die von Manzey (2008) verwendete klare Abgrenzung der beiden Begriffe übernommen. Er bezeichnet mit dem Begriff „Automatisierung“ den Prozess und mit dem Begriff „Automation“ das Ergebnis. Die Definition der Automatisierung als Prozess wird von Hacker (1998) geteilt.

Im Allgemeinen wird unter Automation die Unterstützung oder Ausführung von Aufgaben durch technische Systeme verstanden, welche vormals allein durch den Menschen bewältigt wurden (Manzey, 2008; Parasuraman, 2000; Parasuraman & Riley, 1997; Parasuraman et al., 2000; Sheridan, 2000, 2002; Wandke & Wetzenstein, 2004; Wickens et al., 2004; Wiczorek & Zinn, 2008)<sup>74</sup>. Interessanterweise werden im Alltag gegenwärtige oder zukünftige technische Errungenschaften mit Automatisierung verbunden. Bereits seit langem existierende Automa-

---

<sup>73</sup> Der Nutznießer einer Automation ist nicht allein ein Operateur bei seiner Arbeitstätigkeit, bei dem es sich zumeist um einen sehr gut ausgebildeten Spezialisten handelt. Ebenso kann es sich um wenig bis gar nicht technikerfahrene oder -affine Menschen handeln (Wandke & Wetzenstein, 2004), welche eine Automation auch ohne spezielle Ausbildung privat oder beruflich benutzen. Daher wird in der vorliegenden Abhandlung bei der Darstellung des theoretischen Hintergrunds vom Menschen oder Benutzer allgemein und nicht nur im Speziellen vom Operateur gesprochen. Die dargestellten Untersuchungsergebnisse wurden zumeist an naiven Personen, welche in Laborstudien als Operateure fungierten, oder realen Operateuren bei der Überwachung und Führung von dynamischen Systemprozessen gewonnen. Diese Ergebnisse weisen eine Allgemeingültigkeit für den Menschen auf.

<sup>74</sup> Bei einigen der Autoren findet sich anstelle der Automation noch der Begriff der Automatisierung.

tionen werden jedoch als solche oftmals nicht mehr vom Systembenutzer wahrgenommen (Hauß & Timpe, 2000; Winner, Hakuli, & Wolf, 2009), etwa der elektrische Anlasser anstelle eines Kickstarters beim Motorrad. Ebenso können durch eine Automatisierung gänzlich neue Aufgaben umgesetzt werden (Hauß & Timpe, 2000). Hierbei kann es sich um Aufgaben handeln, zu denen der Mensch aufgrund seiner physischen, psychischen und kognitiven Fähigkeiten nicht oder nur unzureichend in der Lage ist (Gründl, 2005; Hauß & Timpe, 2000; Parasuraman, 2000; Parasuraman et al., 2000; Seifert, 1994; Wandke & Wetzenstein, 2004). Das gesamte MMS kann dadurch eine höhere Leistung erbringen (Seifert, 1994). Außerdem ermöglicht eine solche Aufgabenerweiterung dem Benutzer des automatisierten Systems neue Wege zur Zielerreichung (Hauß & Timpe, 2000).

Eine Automation unterstützt im Allgemeinen den Menschen (Hauß & Timpe, 2000; Manzey, 2008). Sie dient insbesondere dazu, seine Tätigkeit:

- zu vereinfachen (Maier, 2014; Sheridan, 2002; Wandke & Wetzenstein, 2004),
- zeitlich zu verkürzen (Krinner, 2009; Sheridan, 2000, 2002),
- effizienter (Hauß & Timpe, 2000),
- produktiver (Kaber, Onal, & Endsley, 2000; Sarter et al., 1997),
- verlässlicher (Manzey, 2008; Wiczorek & Zinn, 2008),
- sicherer (Gründl, 2005; Hauß & Timpe, 2000; Manzey, 2008; Sarter et al., 1997; Wandke & Wetzenstein, 2004; Wiczorek & Zinn, 2008) und
- komfortabler zu machen (Wandke & Wetzenstein, 2004).

Durch eine Automation soll eine optimale Systemleistung erreicht werden (Seifert, 1994), indem der Benutzer physisch und mental entlastet wird (Cummings & Thornburg, 2011; Maier, 2014; Manzey, 2008; Sarter et al., 1997; Wiczorek & Zinn, 2008), ohne ihn jedoch zu unterfordern (Gründl, 2005). Dadurch können Fehler (Cummings & Thornburg, 2011; Seifert, 1994; Sheridan, 2000, 2002) und sich aus Fehlern ergebende Unfälle vermieden bzw. zumindest deren Auswirkungen vermindert werden (Gründl, 2005). Außerdem soll der Mensch weniger Training für den Umgang mit einem automatisierten technischen System benötigen (Sarter et al., 1997).

Eine Aufgabe wird als Funktion bezeichnet, wenn diese von einem technischen System ausgeführt wird (Seifert, 1994). Das automatisierte System kann Aufgaben aus allen Phasen des menschlichen Informationsverarbeitungsprozesses ausführen, wobei diese Funktionen die unterschiedlichsten Merkmale aufweisen (Wandke & Wetzenstein, 2004). Somit wird die Umsetzung einer Automation durch die auszuführende Aufgabe sowie ggf. durch die Bedürfnisse und Fähigkeiten des Menschen bestimmt. Dabei spielen ergonomische respektive gebrauch-

taugliche (Vöhringer-Kuhnt, 2008) wie auch ethische (Hauß & Timpe, 2000) Aspekte eine entscheidende Rolle. Ebenso müssen technische, wirtschaftliche (Hauß & Timpe, 2000), marketingstrategische sowie rechtliche Bedingungen (Vöhringer-Kuhnt, 2008) berücksichtigt werden. Letzteres betrifft u. a. die Gewährleistung der Funktionssicherheit des Systems (Vöhringer-Kuhnt, 2008) und die Systembeherrschbarkeit durch den Menschen (J. Schwarz, 2007).

#### 4.1.2 Probleme durch Automation

Durch eine Automatisierung von Aufgaben verändert sich die Tätigkeit des Systembenutzers und damit dessen Rolle im MMS (Hauß & Timpe, 2000; Maier, 2014; Manzey, 2008; Parasuraman et al., 2000; Sarodnick & Brau, 2006; Sarter et al., 1997; Wiczorek & Zinn, 2008). Damit gehen jedoch nicht nur die intendierten Vorteile einher. So sind auf Seiten des Benutzers eines automatisierten Systems vor allem ein häufig unangemessenes Systemvertrauen, ein sinkendes Situationsbewusstsein und sogar Fertigkeitsverluste zu beobachten. Nachfolgend werden diese drei Problembereiche näher dargestellt.

##### 4.1.2.1 Unangemessenes Vertrauen in die Automation

**Missbrauch:** Ein unangemessenes Vertrauen des Benutzers in ein automatisiertes System kann zu einem nicht spezifikationsgemäßen Gebrauch von diesem führen (J. Schwarz, 2007). Aufgrund dieses Systemmissbrauchs kann es zu sicherheitskritischen Situationen kommen.

**Misstrauen (engl. Mistrust):** Ebenso kann der Benutzer gegenüber einem automatisierten System ein Misstrauen entwickeln, selbst wenn dieses eine hohe Zuverlässigkeit aufweist (Manzey, 2008; Rocholl, 2010). In Folge dessen, kann es zu mangelnder Akzeptanz (Hauß & Timpe, 2000), mangelnder Nutzung (engl. Disuse, s. Parasuraman & Riley, 1997) oder gar Ablehnung (Lindberg, 2012) des Systems kommen.

**Übervertrauen (engl. Overtrust/Overreliance):** Ein Systembenutzer kann genauso gut ein Übervertrauen in ein automatisiertes System aufbauen. Vor allem, wenn dessen Vigilanz zu stark reduziert (Lindberg, 2012) oder dieser zu hoch beansprucht ist (Kaber & Endsley, 1997; Kaber & Riley, 1999; Kaber et al., 2009; Manzey, 2008; Sheridan, 2002; Wickens et al., 2004; Wiczorek & Zinn, 2008). Dieses Übervertrauen fällt i. d. R. umso stärker aus, je technikaffiner ein Mensch ist (Manzey, 2008). Daraus wiederum kann sich eine fehlerhafte Verwendung (engl. Misuse, s. Parasuraman & Riley, 1997) der Automation in Form einer mangelnden Überwachung/Nachlässigkeit (engl. Complacency, s. Manzey, 2008; Parasuraman, 2000; Parasuraman, Molloy, & Singh, 1993; Parasuraman et al., 2000; Parasuraman & Wickens,

2008; Sarter et al., 1997; Wiczorek & Zinn, 2008) und Verantwortungsabgabe (Lindberg, 2012) ergeben.

#### **4.1.2.2 Sinkendes Situationsbewusstsein**

Eine unzureichende Systemüberwachung kann beim Benutzer zu einem Verlust seines Situationsbewusstseins führen (Endsley, 1995b; Endsley & Kiris, 1995; Kaber & Endsley, 1997; Kaber et al., 2000; Kaber, Riley, Tan, & Endsley, 2001; Manzey, 2008; Parasuraman & Wickens, 2008). Zudem ist die Wahrscheinlichkeit hoch, dass dem Benutzer eine Veränderung in einem dynamischen System kaum auffällt, wenn diese durch eine Automation ohne seine aktive Beteiligung verursacht wird (Endsley & Kiris, 1995). Dem Benutzer fällt es oftmals schwer, die verschiedenen Aktivitäten eines automatisierten Systems wahrzunehmen (Sarter et al., 1997). Die Verringerung des Situationsbewusstseins wird häufig verstärkt, wenn eine schlecht gestaltete Mensch-Maschine-Schnittstelle mangelhafte Rückmeldungen gibt (Kaber & Endsley, 1997) oder beim Menschen ein unangemessenes mentales Modell oder eine Irritation über den aktuellen Systemmodus (engl. Mode Confusion, s. Lindberg, 2012; Manzey, 2008; Sarter et al., 1997) vorliegen. Dadurch wiederum können Störungen im Systemprozess oder Fehler einer Automation übersehen werden (Manzey, 2008) und die MMS-Leistung kann absinken (Endsley, 1987; Endsley & Kiris, 1995; Kaber & Endsley, 1997; Kaber et al., 2000; Kaber et al., 2001).

#### **4.1.2.3 Fertigkeitsverlust**

Bei einem Systembenutzer kann es durch die fehlende Ausführungsübung einer automatisierten Aufgabe zu einer langsameren Fehlerentdeckungs- und -behebungszeit kommen (Kaber et al., 2009; Wiener & Curry, 1980). Möglich ist dadurch auch der Verlust der notwendigen kognitiven und/oder motorischen Fertigkeiten (engl. Skill Degradation) zur Bewältigung dieser Aufgabe (Kaber & Endsley, 1997, 2004; Kaber & Riley, 1999; Kaber et al., 2009; Lindberg, 2012; Manzey, 2008; Parasuraman et al., 2000; Parasuraman & Wickens, 2008; Sheridan, 2002; Wickens et al., 2004; Wiczorek & Zinn, 2008).

An dieser Stelle sei ein weiteres Problem genannt, wenn auch dieses nicht den Fertigkeitsverlust, sondern den Fertigaufbau betrifft. So kann aufgrund der Neuartigkeit oder Komplexität eines automatisierten Systems beim Benutzer ein erhöhter (Seifert, 1994) oder andersartiger (Sarter et al., 1997) Schulungsbedarf zum Umgang mit diesem bestehen.

Die mit einer Automatisierung einhergehenden Probleme werden umso intensiver, je reliabler (Parasuraman, Molloy et al., 1993) und leistungsfähiger das technische System ist

(Endsley & Kiris, 1995; C. A. Miller & Parasuraman, 2007).

#### **4.1.3 Mensch-Maschine-Funktionsteilung**

Aufgrund dieser vielfältigen Probleme, die mit einer Automation im Zusammenhang stehen, sollte nicht alles, was automatisiert werden kann, auch tatsächlich automatisiert werden. In dem Forschungsgebiet „Mensch-Maschine-Funktionsteilung“ geht es um die Klärung der Frage, wie die Zuordnung der Aufgaben an den Menschen bzw. die Maschine in einem MMS optimal zu gestalten ist (Hauß & Timpe, 2000). Es existieren in der Literatur bereits verschiedene Ansätze zur Bestimmung einer adäquaten Funktionsteilung für ein MMS. Die Funktionsallokation sollte an den Fähigkeiten, Bedürfnissen und Erwartungen des Systembenutzers ausgerichtet sein. Das Ziel ist es, ihn zu unterstützen und nicht zu ersetzen (Hauß & Timpe, 2000).

Jahrzehntelang, bis in die 1980er Jahre hinein (Wiczorek & Zinn, 2008), wurde jedoch ein technik- respektive kostenzentrierter Ansatz bei der Automatisierung eines Systems verfolgt (Manzey, 2008). Die Fähigkeiten und Bedürfnisse des Menschen fanden hierbei keinerlei Berücksichtigung. Heute noch lassen sich Entwicklungsprozesse beobachten, die durch diesen Ansatz geprägt sind. Dabei werden die Funktionen zwischen Benutzer und technischem System nach dem Kriterium der Wirtschaftlichkeit aufgeteilt (Endsley & Kaber, 1999; Hauß & Timpe, 2000; Kaber & Endsley, 1997; Parasuraman et al., 2000). Diesem Ansatz wird auch die maximale Automation zugeschrieben (Hauß & Timpe, 2000). Es wird automatisiert, was technisch möglich ist (Kaber & Endsley, 2004; Sarter et al., 1997). Lediglich die übrig bleibenden Funktionen werden dem Systembenutzer überlassen (Kaber & Endsley, 1997; Kaber et al., 2000; Kaber et al., 2009; Parasuraman et al., 2000). Dabei handelt es sich meist um die Systemüberwachung (Kaber et al., 2000; Kaber et al., 2001; Sarter et al., 1997; Sheridan, 2002; Wiener & Curry, 1980). Eine solche Tätigkeit ist für den Menschen jedoch nicht sonderlich geeignet (Endsley, 1995a; Endsley & Kaber, 1999; Kaber & Endsley, 1997). Er gerät durch diese Form der Funktionsteilung „out-of-the-loop“ des Systemprozesses (Kaber & Endsley, 1997; Manzey, 2008). Dies wiederum kann mit einer Vielzahl an negativen Konsequenzen verbunden sein (s. Abschnitt 4.1.2), welche dann die menschliche Leistung im Normalbetrieb des Systems und insbesondere im Fehlerfall einer Automation negativ beeinflussen können (Endsley & Kiris, 1995; Kaber & Endsley, 1997). Deshalb sollte eine reine Überwachungstätigkeit des Menschen vermieden werden (Wiczorek & Zinn, 2008).

Eine ebenfalls klassische Perspektive auf die Optimierung der Mensch-Maschine-Funktionsteilung stellt der fähigkeitszentrierte Ansatz dar (Kaber et al., 2000; Manzey, 2008). Bei

diesem wird für jede Aufgabe der Überwachung und Führung eines Systems ein Leistungsvergleich zwischen dem Menschen und der Automation vorgenommen (Hauß & Timpe, 2000; Parasuraman et al., 2000). Die Ergebnisse schlagen sich in einer Fähigkeitsliste nieder. Die erste und wohl bekannteste Auflistung dieser Art wurde von Fitts bereits im Jahr 1951 publiziert. Diese ist im deutschsprachigen Raum als MABA-MABA-Liste (kurz für „Men are better at, machines are better at“) und im englischen Raum nach deren Schöpfer Fitts' list benannt. Auf der Basis einer solchen Liste werden der menschlichen und der technischen Komponente eines MMS jeweils die Aufgaben zugewiesen, für welche diese jeweils besser geeignet sind (Manzey, 2008). Jedoch hängt eine solche Zuteilung letztlich von den technischen Fähigkeiten ab (Hauß & Timpe, 2000). Erweist sich die Automation als ungeeigneter für die Ausführung einer Funktion, muss der Mensch diese Aufgabe übernehmen, selbst wenn er dafür nicht befähigt ist. Ein weiterer Nachteil des fähigkeitszentrierten Ansatzes liegt im rapiden technologischen Fortschritt begründet. So können Vergleichsergebnisse, welche heute gewonnen werden, morgen bereits schon veraltet und womöglich nur noch eingeschränkt gültig bis ungültig sein. Permanente Aktualisierungen der Listen wären daher von Nöten (Parasuraman et al., 2000; Sheridan, 2002).

Eine Optimierung der Funktionsteilung zwischen Mensch und Maschine, welche nicht allein auf die technischen und wirtschaftlichen Aspekte (Wandke & Wetzenstein, 2004) oder den Fähigkeitsvergleich der MMS-Komponenten zielt, strebt der menschenzentrierte Ansatz der Automatisierung (engl. human-centered-automation) nach Billings (1997) an. Hier finden die physischen und psychischen Merkmale des Menschen Beachtung (Herczeg, 2009). Als Kriterium wird die „menschengerechte Aufgabe“ verwendet (Hauß & Timpe, 2000). Danach soll eine Aufgabe vom Systembenutzer, dessen Bedürfnisse (Sarter et al., 1997) und Fähigkeiten zu berücksichtigen sind, gut ausgeführt werden können (Kaber & Endsley, 2004). Darüber hinaus muss sie dem Benutzer auch zumutbar sein (Wandke & Wetzenstein, 2004). Außerdem soll durch die Aufgabe eine Fehlbeanspruchung des Menschen vermieden sowie dessen Kompetenzen erhalten und ausgebaut werden (Wandke & Wetzenstein, 2004). Ferner wird im Gegensatz zu den beiden älteren Ansätzen die Leistung eines MMS nicht nur im Normalbetrieb, sondern auch im Fehlerfall der Automation betrachtet (Kaber & Endsley, 1997). Nur durch ein solches Vorgehen kann eine Automatisierung zu einem effektiven und sicheren MMS führen (Parasuraman, 2000).

In den letzten Jahren vollzog sich zudem ein Perspektivwechsel, der den Aspekt der Konkurrenz stärker zu Gunsten der Kooperation zwischen Mensch und Maschine in den Hintergrund treten ließ. Während Manzey (2008) diesen Perspektivwandel noch dem menschenzen-



trierten Ansatz zuordnet, erblicken Hauß und Timpe (2000) darin bereits einen neuen, den mensch-maschine-systemzentrierten Ansatz. Hierbei stehen das Gesamtsystemziel als auch die Ressourcen von Mensch und Maschine im Fokus (Grote, Weik, Wäfler, Zölch, & Ryser, 1999). Der Mensch und die Maschine werden als voneinander abhängige (Manzey, 2008), gleichberechtigte „Partner“ betrachtet (Christoffersen & Woods, 2002, zitiert nach Wiczorek & Zinn, 2008), welche kooperieren (Sarter et al., 1997), sich gegenseitig koordinieren, kontrollieren, unterstützen und ergänzen (Wiczorek & Zinn, 2008). Ein hohe MMS-Leistung kann nur erreicht werden, wenn das Zusammenspiel zwischen diesen beiden „Partnern“ optimiert wird (Grote et al., 1999). Dafür müssen deren gegenseitige Kommunikations- und Koordinationsbedürfnisse berücksichtigt werden (Sarter et al., 1997). Ein Kriterium dieses Ansatzes stellt die Flexibilität dar (Hauß & Timpe, 2000). Durch den Einsatz einer flexiblen (dynamischen) Automation können Funktionen sowohl vom Menschen als auch vom System ausgeführt werden (s. Abschnitt 4.1.6.1).

Forschung und Praxis haben sich über die Jahrzehnte bis zum heutigen Tag zunehmend von einer exklusiven und fixen Zuordnung einer Aufgabe an die menschliche oder technische Komponente eines Systems entfernt. Verschiedenste Automationsvarianten mit unterschiedlichster Komplexität wurden inzwischen entwickelt, evaluiert und in der Praxis eingesetzt, welche zwischen einer manuellen und vollautomatischen Funktionsausführung liegen (Endsley & Kaber, 1999; Parasuraman et al., 2000; Sheridan, 2002). Dabei kann ein MMS unterschiedlichste Automationsvarianten gleichzeitig innehaben (Hauß & Timpe, 2000; Wandke & Wetzenstein, 2004). Die Automation reicht von einer fixen bis hin zu einer kooperativen Auslegung (Hauß & Timpe, 2000). Ebenso sind häufig mehrere Menschen und Systeme miteinander verbunden (Hauß & Timpe, 2000; Johannsen, 1993; Manzey, 2008; Sarter et al., 1997), wodurch die Funktionsaufteilung noch komplexer wird.

#### **4.1.4 Begriffsbestimmung „Automatik“ und „Assistenzsystem“**

Eine Automation kann verschiedene Ausmaße annehmen, welche ein Kontinuum mit den Polen „manuelles“ und „vollautomatisches“ System aufspannen. Im Rahmen dieser Arbeit wird ein vollautomatisches System als Automatik bezeichnet. Die jeweiligen Stufen zwischen den beiden Polen stellen nach Wandke und Wetzenstein (2004) Assistenzsysteme dar. Dabei sind die Übergänge fließend (Hauß & Timpe, 2000). Diese Sicht wird von der Autorin geteilt. Allerdings existiert auch für den Begriff „Assistenzsystem“ keine einheitliche Definition. So beschreiben Hauß und Timpe (2000) diese Zwischenstufen als Unterstützungssysteme und lediglich die Unterstützungssysteme, welche nicht autark Entscheidungen treffen, Befehle geben oder Handlungen ausführen als Assistenzsysteme. Nach dieser Definition behält der

Systembenutzer bei der Interaktion mit einem Assistenzsystem jederzeit die Kontrolle. Dagegen können nach der von Wandke und Wetzenstein (2004) vertretenen Begriffsbestimmung Assistenzsysteme ebenso die Kontrolle innehaben. Jedoch bleibt es immer Aufgabe des Assistenzsystems den Menschen bei seiner Zielerreichung zu unterstützen (Hauß & Timpe, 2000).

Ein Assistenzsystem entspricht zudem der Interaktionsform zwischen dem Menschen und dem technischen System. Somit zeichnet sich ein Assistenzsystem zum einen durch seine Interaktivität, zum anderen aber auch durch den Zugang aus, den es dem Menschen zu den Funktionen eines technischen Systems anbietet bzw. erleichtert. (Wandke & Wetzenstein, 2004)

#### **4.1.5 Automatisierungsstufen**

Für das Ausmaß der Automation eines Systems existieren zwei Operationalisierungen: der Automatisierungsgrad und die Automatisierungsstufe (Hauß & Timpe, 2000).

Der Automatisierungsgrad quantifiziert (Hauß & Timpe, 2000) durch den prozentualen Anteil der automatisierten Funktionen eines MMS das Ausmaß der Automation (Johannsen, 1993). Dieses Vorgehen ist jedoch nur bedingt geeignet zur Darstellung des Automationsausmaßes eines MMS, da die Anzahl der automatisierten Funktionen in Abhängigkeit von der gewählten Granularität der Systembetrachtung variiert (Hauß & Timpe, 2000).

Die Automatisierungsstufen stellen eine eher qualitative Klassifikation dar (Hauß & Timpe, 2000). Dabei handelt es sich um eine überblicksartige Darstellung der verschiedenen Automationsmöglichkeiten, die nach deren Kernmerkmalen strukturiert sind. Durch diese Klassifikation soll die Komplexität der Automationsvarianten reduziert, geordnet und übersichtlicher gemacht werden. Sie bietet Entwicklern von MMS eine theoretische und teilweise experimentell getestete Basis für die Identifizierung und Gestaltung einer für den jeweiligen Einsatzbereich angemessenen Funktionsteilung zwischen Mensch und Maschine (Endsley & Kaber, 1999; Parasuraman et al., 2000; Sheridan, 2000, 2002; Sheridan & Verplank, 1978; Wandke, 2005; Wandke & Wetzenstein, 2004). Weil jedes MMS speziell ist und individuelle Anforderungen an die Funktionsteilung stellt, kann ein solches Klassifikationssystem jedoch keine allgemeingültigen Musterlösungen vorgeben (Sheridan, 2002). Um eine geeignete Funktionsteilung für ein bestimmtes MMS zu finden, ist eine Evaluation verschiedener Varianten der Funktionsteilung unter Betrachtung der MMS-Leistung im Normalbetrieb als auch im Fehlerfall der Automation notwendig (Kaber et al., 2009; Parasuraman et al. 2000; Rocholl, 2010). Die Automatisierungsstufen sind in unterschiedlichsten Modellen oder Taxonomien beschrieben (s. bspw. Endsley, 1987, 1997; Endsley & Kaber, 1999; Hollister, 1986;

Parasuraman et al., 2000; Sheridan, 1988, 1992, 2000; Sheridan & Verplank, 1978; Timpe, 1998; Wandke & Wetzenstein, 2004). Nachfolgend werden einige der bekanntesten Klassifikationen von Automationsvarianten für die Prozessüberwachung und -führung kurz vorgestellt.

Einer der frühesten Versuche zur Klassifizierung von Automationsvarianten ist die Taxonomie von Sheridan und Verplank (1978). Sie gliedern das Kontinuum von einem manuellen bis zu einem vollautomatischen System<sup>75</sup> für die Entscheidungsfindung hinsichtlich eines Eingriffs in einen selbstständig ablaufenden Prozess (Wandke & Wetzenstein, 2004) mit anschließender Aktionsausführung durch den Menschen oder das System in zehn<sup>76</sup> Stufen auf. Von Stufe zu Stufe wird dabei das technische System autonomer. Diese Taxonomie ist jedoch nicht geeignet, um der Fülle der mittlerweile existierenden Automationen gerecht zu werden. Neben der Entscheidungsfindung und Handlungsausführung sind weitere Handlungsphasen relevant für die Kategorisierung von Automationsvarianten (Wandke & Wetzenstein, 2004). Diese Handlungsphasen lassen sich aus den zyklischen Handlungsregulationstheorien ableiten, welche die menschliche Tätigkeit zur zielgerichteten Bewältigung einer Aufgabe in verschiedene hierarchisch-sequentielle Stufen oder Phasen unterteilen (Dörner, 1989; G. A. Miller, Galanter, & Pribram, 1960; Norman, 1986, 1988; Rasmussen, 1986; Volpert, 1987). Automationen können den Menschen in allen Handlungsphasen unterstützen oder gar ersetzen. Aus diesem Grund wurden Taxonomien entwickelt, welche sich explizit auf die menschlichen Handlungsphasen beziehen.

Eine dieser Taxonomien stammt von Endsley<sup>77</sup> (1997), die sie später gemeinsam mit Kaber (1999) veröffentlichte. In dieser Taxonomie werden die menschlichen Handlungsphasen durch die vier Funktionskategorien „Überwachung“, „Strategieentwicklung“, „Strategieauswahl“ und „Umsetzung“ abgebildet. In Abhängigkeit davon, ob eine Funktion in diesen vier Kategorien durch den Menschen und/oder das technische System ausgeführt wird, ergeben sich 81 verschiedene Automatisierungsstufen. Endsley (1997) ordnet hiervon zehn praxisrelevante Stufen in ihre Taxonomie. Damit bietet diese Klassifikation jedoch nur drei verschiedene Ausführungsmöglichkeiten in der jeweiligen menschlichen Handlungsphase an (Hauß & Timpe, 2000).

Eine weitere Taxonomie mit Bezug auf die menschlichen Handlungsphasen stammt von Parasuraman et al. (2000). Dieses Klassifikationssystem wird nachfolgend etwas ausführlicher

---

<sup>75</sup> Die Automatisierungsstufen von Sheridan und Verplank (1978) beziehen sich auf ein Mensch-Computer-System. Diese sind jedoch ohne Probleme auf alle Mensch-Maschine-Systeme übertragbar.

<sup>76</sup> Später wurden von Sheridan (1988, 1992, 2000, 2002) verschiedene Versionen mit unterschiedlicher Stufenanzahl publiziert.

<sup>77</sup> Die Automatisierungsstufen von Endsley (1997) beziehen sich ebenfalls auf ein Mensch-Computer-System. Auch sie sind problemlos auf alle Mensch-Maschine-Systeme übertragbar.

vorgestellt, da dies für die Entwicklung eines Instruments zur Bewertung der Güte von Automatikkonzepten im Rahmen dieser Dissertation eine besondere Rolle spielte (s. Abschnitt 5.2.2). Diese Taxonomie unterscheidet die vier Funktionsklassen: „Informationsakquise“, „Informationsanalyse“, „Entscheidungsfindung und Handlungsauswahl“ sowie „Handlungsausführung“ (s. Abbildung 20; Parasuraman et al., 2000).

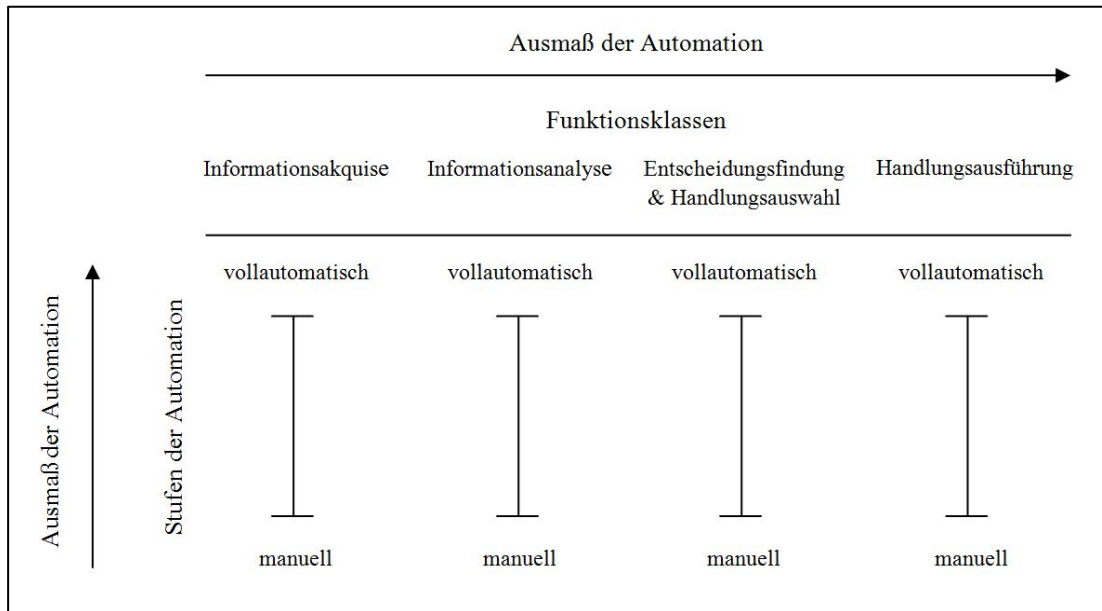


Abbildung 20: Funktionsklassen des Modells „Arten und Stufen von Automation“ nach Parasuraman et al. (2000) und das Ausmaß der Automation nach Onnasch, Wickens, Li und Manzey (2014)

In der Phase der „Informationsakquise“ finden sensorische Prozesse zur selektiven Erfassung von verschiedensten Umwelt- und Systemdaten statt. Diese Phase kann bspw. durch technische Sensoren unterstützt werden, welche für den Menschen schwach wahrnehmbare oder unterschwellige Signale messen. Dabei können dem Menschen die erfassten Rohdaten unsortiert, bzw. nach bestimmten Kriterien (etwa nach der Priorität) geordnet, hervorgehoben oder bspw. einige gefiltert dargeboten werden. (Parasuraman, 2000; Parasuraman et al., 2000) Ein Beispiel aus dem Kraftfahrzeugbereich sind Nachtsichtassistenten. Diese bieten dem Fahrer eine erweiterte Sicht in der Dunkelheit. Dafür wird die, von einer Wärmebildkamera aufgezeichnete, abgestrahlte Wärme der Umgebungsobjekte dem Fahrer in Schwarzweißbildern angezeigt.

Die Informationsanalyse umfasst bereits höhere kognitive Prozesse, die z. B. die Verarbeitung und Interpretation der eingehenden Signale beinhaltet. Automationen dieser Phase unterstützen die Auswertung der erfassten Umwelt- und Systemdaten mit dem Ziel einer Bewertung oder Antizipation der Umwelt- und Systemsituation. Dabei lassen sich Umwelt oder Systemdaten zu einer neuen Variable integrieren, um so die Wahrnehmung des System-

benutzers zu erweitern. (Parasuraman, 2000; Parasuraman et al., 2000)

Beispielhaft für Automationen dieser Phase sind Alarmsysteme, die vor bestimmten Umwelt- oder Systemzuständen warnen (Manzey, 2008). Ein konkretes Beispiel, erneut aus dem Kraftfahrzeugbereich, ist die visuelle Reifendruckwarnung durch ein Reifendruck-Kontrollsystem. Sobald ein unangemessener Reifendruck erreicht ist, wird der Fahrer gewarnt.

In der dritten Phase unterstützen Automationen die Entscheidungsfindung und Handlungsauswahl des Menschen (Parasuraman, 2000). Das kann bspw. durch Handlungsempfehlungen geschehen (Manzey, 2008; Parasuraman et al., 2000). Hierfür exemplarisch ist das, in Flugzeugen eingesetzte, Traffic Collision Avoidance System (TCAS), welches bei drohender Kollision von zwei Flugzeugen den Piloten eine kurze verbale Ausweichempfehlung liefert. Die verschiedenen Stufen der Automationen dieser Phase werden gut durch die Taxonomie von Sheridan und Verplank (1978) dargestellt (Parasuraman, 2000; Parasuraman et al., 2000).

Die vierte Phase umfasst Automationen, welche die Handlungsausführung unterstützen oder sogar vollständig übernehmen (Parasuraman, 2000; Parasuraman et al., 2000). Beispielhaft aus dem Bereich der Luftfahrt ist der Autopilot, welcher vom Piloten oder dem Flight Management System vorgegebene Sollwerte für den Kurs, die Höhe und Geschwindigkeit des Flugzeugs umsetzt (Liebig, 1994).

In jeder dieser vier Funktionsklassen kann das Ausmaß der Automation auf einem Kontinuum von manuell bis vollautomatisch variieren (Parasuraman, 2000; Parasuraman et al., 2000), wobei die einzelnen Stufen der Automation nicht explizit definiert sind (Kaber & Endsley, 2004). Somit weist diese Taxonomie eine höhere Differenzierung der Automationsmöglichkeiten je Funktionsklasse auf als jene von Endsley (1997). In einem MMS können in den Funktionsklassen Automationen verschiedenster Stufen zum Einsatz kommen (Parasuraman, 2000; Parasuraman et al., 2000). Nach Onnasch, Wickens, Li und Manzey (2014) steigt das Ausmaß der Automation an, je höher die Stufe der Automation und je weiter hinten im Informationsverarbeitungsprozess die unterstützte Phase liegt.

Abschließend soll hier die Taxonomie von Wandke (2005) und Wetzstein (2004) kurz vorgestellt werden, da diese ebenfalls relevant für das ATEO-Projekt war. Diese klassifiziert Assistenzsysteme nicht nur für die Prozessüberwachung und -führung von Systemen, sondern allgemein für die Unterstützung von Mensch-Technik-Interaktionen. Dafür werden fünf Dimensionen verwendet. Eine Dimension, die „Handlungsphase“ umfasst neben den Phasen „Informationsaufnahme“, „Informationsanalyse und -integration“, „Entscheidung“ sowie „Aktionsausführung“ noch die beiden Phasen „Motiv- und Zielbildung“ sowie „Effektkontrolle“. Die zweite Dimension bildet die „Art der Assistenz“ ab, welche abhängig von der

zu unterstützenden Handlungsphase ist. Dabei sind die Assistenzarten der Phase der „Entscheidung und Aktionsauswahl“ wiederum an das Modell von Sheridan und Verplank (1978) angelehnt. Die Automation eines MMS kann aus mehreren solcher Assistenzarten bestehen. Dadurch können wiederum mehrere Handlungsphasen oder eine Handlungsphase durch mehrere Assistenzstufen unterstützt werden. Die weiteren drei Dimensionen umfassen die „Anpassbarkeit des Assistenzsystems“ (fixiert, pre-konfigurierbar, adaptierbar und adaptiv), die „Art der Initiative des Assistenzsystems“ (aktiv und passiv) und die „Modalität der Informationsvermittlung“ zwischen Automation und Mensch (monomodal, multimodal und amodal). Durch diese fünf Dimensionen bietet diese Taxonomie trotz der breiten Klassifikation der Assistenzsysteme eine endliche Anzahl an definierenden Elementen. Somit wird eine genaue Unterscheidung zwischen den vielfältigen Assistenzsystemen möglich, ohne jedoch den geschlossenen und überblicksartigen Charakter einer Taxonomie zu verlieren.

#### **4.1.6 Optimales Ausmaß der Automation**

Für die einzelnen menschlichen Handlungsphasen wurden in vielfachen empirischen Überprüfungen optimale Ausmaße der Automation in Abhängigkeit des Systems, der Aufgabe sowie des Einsatzgebietes identifiziert (Clamann, Wright, & Kaber, 2002; Endsley, 1987; Endsley & Kiris, 1995; Kaber et al., 2000; Parasuraman et al., 2000; Sharit, 1984). Dabei zeigt sich eindeutig im Normalbetrieb eines Systems ein Anstieg der MMS-Leistung, je höher das Ausmaß der Automation ist (Endsley & Kaber, 1999; Kaber et al., 2000; Onnasch et al., 2014). Ebenso, wenn auch nicht ganz so eindeutig, findet sich eine Reduzierung der physischen und mentalen Beanspruchung des Menschen, umso höher das Ausmaß der Automation ist (Endsley, 1995b; Kaber & Endsley, 2004; Kaber et al., 2000; Kaber et al., 2001; Onnasch et al., 2014). Jedoch kann diese Reduzierung über ein angemessenes Beanspruchungsniveau hinaus zu einer Unterforderung des Menschen führen und damit einhergehend zu einem zu starken Vigilanzabfall (Kaber & Endsley, 2004; Wiczorek & Zinn, 2008). In Folge dessen, ist wiederum ein Verlust des Situationsbewusstseins nicht auszuschließen (Endsley, 1995b). Andererseits können selbst bei einem hohen Ausmaß an Automation für den Menschen viele Eingabe- und Überwachungsaufgaben verbleiben oder neu hinzukommen, welche dessen physische oder mentale Beanspruchung aufrechterhalten, lediglich umverteilen oder gar erhöhen können (Bainbridge, 1983; Kaber & Endsley, 2004; Kaber et al., 2001; Parasuraman, 2000; Parasuraman et al., 2000; Sarter et al., 1997; Wiczorek & Zinn, 2008). Des Weiteren gehen mit einem höheren Ausmaß der Automation ein größerer Nachlässigkeits-Effekt, ein reduzierteres Situationsbewusstsein (Endsley, 1995b; Endsley & Kaber, 1999; Kaber et al., 2000; Onnasch et al., 2014) sowie ein wahrscheinlicher Verlust der Fertigkeiten des Menschen einher (Endsley

& Kiris, 1995; Kaber, Wright, Prinzel III, & Clamann, 2005; Parasuraman, 2000; Parasuraman et al., 2000). Dadurch sinkt im Fehlerfall der Automation die gesamte MMS-Leistung stärker ab (Endsley & Kaber, 1999; Endsley & Kiris, 1995; Kaber & Endsley, 1997; Kaber et al., 2000; Onnasch et al., 2014; Parasuraman, 2000; Sharit, 1984).

Die mit einer hohen Automation verbundenen Probleme scheinen hauptsächlich durch hohe Stufen der Automation der beiden späteren Phasen des menschlichen Informationsverarbeitungsprozesses, also der Entscheidungsfindung (Manzey, 2008) und der Handlungsausführung, zu entstehen (Onnasch et al., 2014). Daher sollten in diesen beiden Phasen nur für Situationen mit einem geringen Risiko (Parasuraman, 2000) oder hoher zeitlicher Kritikalität (Parasuraman et al. 2000) hohe Stufen der Automation verwendet werden.

Wenn die Entscheidungsfindung durch eine Automation durchgeführt wird, dann verliert der Mensch mangels Übung im Laufe der Zeit seine Fertigkeiten zur Erledigung dieser Aufgabe (Kaber & Riley, 1999; Parasuraman et al., 2000). Das geschieht für kognitive Fertigkeiten schneller als für motorische (Patrick, 1992, zitiert nach Manzey, 2008). Der Systembenutzer benötigt diese Fertigkeiten jedoch, um als „Rückfalllösung“ für eine Automation fungieren zu können (Parasuraman, 2000). Um kognitive Fertigkeiten aufrecht zu erhalten, sollte der Mensch zumindest die automatisierten Systemprozesse kognitiv nachvollziehen können (Lorenz, DiNocera, Röttger, & Parasuraman, 2002, zitiert nach Manzey, 2008). Noch vorteilhafter ist die Entscheidungsfindung durch den Systembenutzer selbst, da dieser dadurch ein ausreichendes Situationsbewusstsein für den Fehlerfall der Automation aufbauen kann (Onnasch et al., 2014). Dabei kann eine Automation durch Vorschläge unterstützend mitwirken (Parasuraman, 2000; Parasuraman et al., 2000). Wenn allerdings eine hohe Stufe der Automation für die Entscheidungsfindung gewählt wurde, sollte nur eine niedrige Stufe der Handlungsausführung vorgesehen werden, um das Involvement des Menschen im Systemprozess und somit sein Situationsbewusstsein aufrechtzuerhalten (Parasuraman et al., 2000). Dies ermöglicht dem Systembenutzer, mehr Fehler einer Automation bei der Entscheidungsfindung zu erkennen und zu korrigieren (Parasuraman et al., 2000). Andererseits sollte bei der Wahl einer hohen Stufe der Automation für die Handlungsausführung keine hohe Stufe für die Entscheidungsfindung gewählt werden (Parasuraman et al., 2000). Dabei ist zu beachten, dass eine vollautomatische Handlungsausführung im Fehlerfall der Automation nicht die beste menschliche Leistung bei der Prozessführung mit sich bringt (Endsley & Kaber, 1999). So weisen nur die niedrigen bis mittleren Stufen der Automation der Handlungsausführung eine wesentlich bessere menschliche Gesamtleistung auf im Vergleich zu einer rein manuellen Handlungsausführung (Kaber & Endsley, 1997).

Dagegen können bei einem reliablen System hohe Stufen der Automation in den Phasen der „Informationsakquise“ sowie der „Informationsanalyse“ Leistungsvorteile mit sich bringen (Galster, Bolia, & Parasuraman, 2002). So kann eine hohe Automation der Informationsakquise in einem menschlichen Team den Bedarf nach Kommunikation reduzieren (Wright & Kaber, 2005). Eine hohe Automation der Informationsanalyse führt zu einer verbesserten Koordination und effektiverer Tätigkeit eines menschlichen Teams (Wright & Kaber, 2005). Dadurch können diese Informationsautomatationen eine angemessene Beanspruchung der Teammitglieder, aber auch einzelner Systembenutzer bewirken (Parasuraman et al., 2000).

Generell zeigen sich bei einem mittleren im Vergleich zu einem hohen Ausmaß an Automation eine höhere Fähigkeit des Systembenutzers, sich aufgrund eines Automationsfehlers wieder in die Prozessführung zurück zu orientieren (Kaber et al., 2001), verminderte „out-of-the-loop“ Leistungsprobleme (Kaber & Endsley, 1997) und ein höheres Situationsbewusstsein (Endsley & Kaber, 1999; Endsley & Kiris, 1995; Kaber & Endsley, 2004; Kaber, et al. 2000; Kaber et al., 2001). Gleichzeitig weist der Mensch zumeist eine für ihn akzeptable Beanspruchung auf im Vergleich zu einer rein manuellen Aufgabenausführung (Kaber & Endsley, 2004).

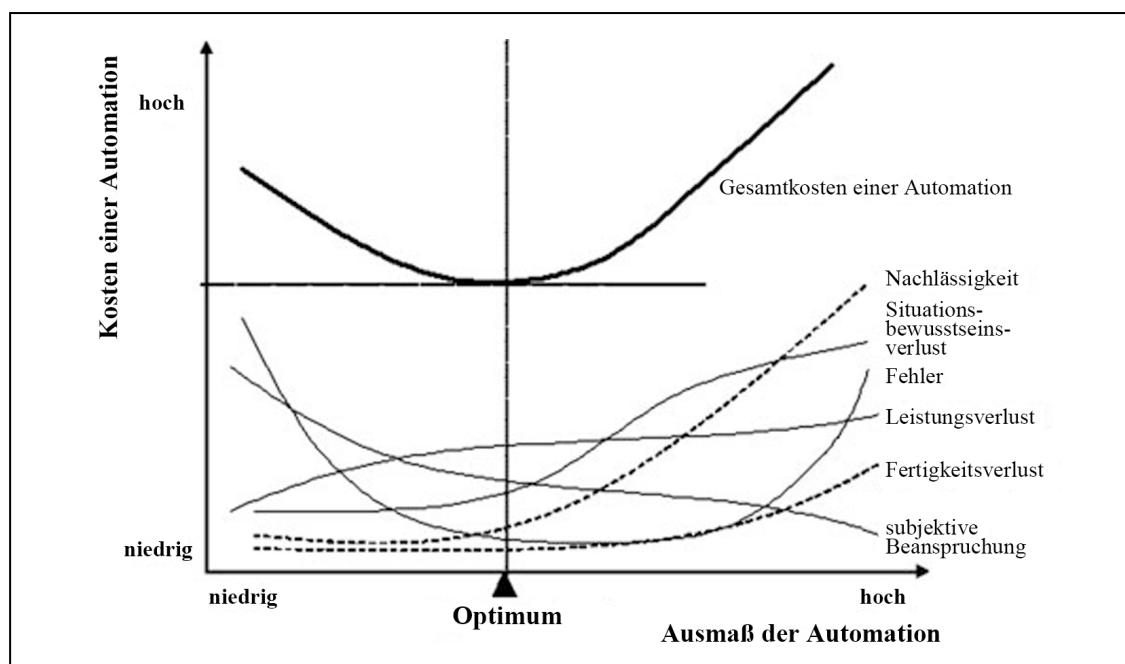


Abbildung 21: U-förmiger Verlauf der Gesamtkosten von Automation nach Kaber et al. (2009, S. 617, ins Deutsche übersetzt durch die Autorin)

Die negativen Auswirkungen der Automation (ermittelt aus Leistung, Fehlern, subjektiver Beanspruchung, Situationsbewusstseinsverlust, Fertigkeitsverlust, Nachlässigkeit) bilden über das Ausmaß der Automation hinweg einen u-förmigen Verlauf mit einem systemabhängigen Minimum im mittleren Bereich (s. Abbildung 21; Kaber et al., 2009). Somit scheint



ein mittleres Ausmaß an Automation am geeignetsten für ein MMS zu sein (Endsley & Kiris, 1995). Dabei bleibt der Mensch bei der Entscheidungsfindung und Handlungsausführung durch die Wahl einer der mittleren Stufen der Automation im Systemprozess involviert (Endsley & Kaber, 1999; Kaber & Endsley, 2004; Onnasch et al., 2014; Parasuraman, 2000; Parasuraman et al., 2000).

#### 4.1.6.1 Flexible Automation

Die Fertigkeiten, das Situationsbewusstsein sowie die physische und mentale Beanspruchung eines Menschen sind bei der Interaktion mit einem technischen, dynamischen System leistungs- (Endsley, 1995, zitiert nach Endsley, Hoffman, Kaber, & Roth, 2007; Kaber et al., 2001) und sicherheitsrelevant (Kaber & Endsley, 1998; Parasuraman et al., 2000; Vöhringer-Kuhnt, 2008, 2010). Dabei bestimmt das Ausmaß einer fixen Automation die Höhe der MMS-Leistung, des Situationsbewusstseins (Kaber & Endsley, 2004) sowie der physischen und mentalen Beanspruchung des Menschen. Jedoch liegt deren jeweilige optimale Ausprägung bei unterschiedlichen Ausmaßen der Automation (s. Abbildung 21). Deshalb scheint der geeignetere Weg zur Berücksichtigung des Menschen, seiner Fähigkeiten, Fertigkeiten und Bedürfnisse in komplexen, technischen Systemen der Einsatz einer flexiblen oder gar kooperativ<sup>78</sup> aushandelnden Automation zu sein. Hierbei variiert die Stufe der Automation von manuell bis hin zu vollautomatisch über den Zeitverlauf in Abhängigkeit von der jeweiligen Situation (Parasuraman, et al. 2000; Scerbo, 1996), der System- oder Benutzerleistung (Parasuraman, Mouloua, & Molloy, 1996), psychophysiologischer Werte (Kaber et al., 2005) bzw. der Selbstentscheidung des Menschen (Hilburn, Molloy, Wong, & Parasuraman, 1993; Kaber & Riley, 1999; Manzey, 2008; Parasuraman, 2000; Parasuraman & Wickens, 2008; Sarter, et al. 1997). Eine Zuordnung der Funktionen durch das technische System wird als adaptive Automation und durch den Menschen als adaptierbare Automation bezeichnet (Manzey, 2008; Wiczorek & Zinn, 2008). Die Zuteilung wird zumeist vom Menschen vorgenommen und nur wenn dieser stark belastet ist, dem technischen System überlassen (Hauß & Timpe, 2000). Allerdings haben Studien gezeigt, dass der Systembenutzer nicht der bessere Entscheider für eine angemessene Funktionsteilung ist (Hilburn et al., 1993; Kaber & Riley, 1999; Scerbo, 1996). Er nimmt häufig einen Wechsel zwischen manueller und automatischer Aufgabenausführung vor, da er subjektiv eine steigende Beanspruchung wahrnimmt, gleichwohl objektiv keine Veränderung stattfindet (Parasuraman, 1997). Des Weiteren kann die Ent-

---

<sup>78</sup> Da die kooperative Automation im Rahmen der dritten ATEO-Projektphase im Fokus der Forschung liegt, wird auf diese Automationsform hier nicht näher eingegangen.

scheidung über die Funktionsteilung zu viele kognitive Ressourcen des Systembenutzers beanspruchen (Kaber & Riley, 1999; Kaber et al., 2001). Auf der anderen Seite geht eine adaptive Automation mit einem Verlust an Situationskontrolle für den Menschen einher, was zu Akzeptanzproblemen des Systems (Manzey, 2008) und Frustration beim Benutzer (Kaber & Riley, 1999) führen kann. Außerdem kann der Mensch von einer Automationshandlung überrascht werden (engl. Automation Surprises), wenn ein vorhergehender automatischer Automationsstufenwechsel unzureichend transparent war (Sarter et al., 1997).

Experimentell konnte belegt werden, dass eine flexible Automation eine höhere Leistung des Menschen verursacht als eine rein manuelle Systemüberwachung (Hilburn et al., 1993) und -führung, vor allem bei der Handlungsausführung (Kaber et al., 2005). Ebenso überwachen Menschen eine Automation stärker, wenn es sich um eine flexible im Vergleich zu einer fixen Automation handelt (Parasuraman et al., 1996). Dadurch erreicht der Benutzer ebenfalls eine höhere Gesamtsystemleistung (Kaber & Endsley, 2004; Parasuraman et al., 1996; Parasuraman, Mouloua, Molloy, & Hilburn, 1993). Zugleich führt eine flexible Automation zu einem höheren Situationsbewusstsein (Kaber & Endsley, 2004) und einer geringeren mentalen Beanspruchung des Menschen im Vergleich zu einer rein manuellen oder generell automatisierten Aufgabenausführung (Hilburn, Jorna, Byrne, & Parasuraman, 1997). Das Situationsbewusstsein und die Beanspruchung des Menschen erreichen dabei ein adäquates Niveau (Hilburn et al., 1997; Manzey, 2008).

Allerdings sollte die Aktivität einer flexiblen Automation nicht zu lange andauern, weil sich andernfalls das Situationsbewusstsein des Menschen verschlechtern und „out-of-the-loop“-Leistungsprobleme auftreten können (Hilburn et al., 1993; Kaber & Endsley, 2004).

Vor allem unter der Bedingung eines erhöhten Zeitdrucks, fällt die Reaktionszeit und Genauigkeit eines Menschen zu gering aus, weshalb in solchen Situationen hohe Stufen einer flexiblen Automation für die Systemführung eingesetzt werden sollten (Moray, Inagaki, & Itoh, 2000).

## **4.2 Entwicklung von Mensch-Maschine-Systemen**

Die Entlastung und Fehlerminimierung eines Menschen durch eine Automation stellt keine Eliminierung von menschlichen Schwächen dar, sondern eine Verlagerung dieser vom Systembenutzer zum -entwickler (Rocholl, 2010; Sheridan & Parasuraman, 2006; Wiczorek & Zinn, 2008). Diese müssen oft unter einengendem Zeit-, Kosten- (Ehrlenspiel, 2003) und Innovationsdruck (Brown & Eisenhardt, 1995; Sarter et al., 1997; VDI, 1993) nutzerfreundliche, beherrschbare, störungsfreie und sichere Systeme erschaffen (Rocholl, 2010). Dieser Druck wird u. a. durch immer kürzere Produktlebenszyklen (Brown & Eisenhardt, 1995;

Verworn, 2005), dem wachsenden Konkurrenzkampf mit Wettbewerbern (VDI, 1993), den technischen Möglichkeiten und aufgrund der hohen Erwartungen der Vorgesetzten bzw. in erster Linie der Systembenutzer bedingt.

In diesem Abschnitt werden der Entwickler von technischen Systemen sowie seine Tätigkeit näher vorgestellt. Dabei wird der Bezug zum Lösen von komplexen Problemen hergestellt.

#### **4.2.1 Entwickler von technischen Systemen**

Bei der Entwicklung von MMS gehört es zu den Aufgaben der Entwickler, eine adäquate Funktionsteilung umzusetzen (Fuld, 1993; Hauß & Timpe, 2000). Dafür muss der Entwickler in der Lage sein, vor auszudenken (Sachse & Furtner, 2009), das zukünftige System mit seinen Funktionen und deren Wirkungen (Hacker & Sachse, 2006) sowie die Systemprozesse und deren Störungen zu antizipieren (Sarodnick & Brau, 2006), um die geeignetste Lösung aus der Menge an Optionen herauszufiltern. Dabei reicht die isolierte Betrachtung des Systems nicht aus. Vielmehr muss prospektiv das Systemverhalten und dessen Auswirkungen in der Einsatzumgebung eingeschätzt werden (Hauß & Timpe, 2000). Dafür muss der Entwickler einerseits große Mengen an Wissen gleichzeitig im Gedächtnis aktivieren, zum anderen aber auch neues Wissen erarbeiten und integrieren. Zudem ist eine adäquate Wissensstruktur aufzubauen, um die relevanten Informationen schnell abrufen zu können (Hacker, 1998; Kintsch & Ericsson, 1996). Vorhandenes Wissen, Mittel und Vorgehensweisen können nicht nur routinemäßig angewendet, sondern es müssen neuartige Kombinationen dieser sowie gänzlich Neues generiert werden (Schlicksupp, 1998). Dabei erfolgt die Kombination zweckorientiert und in Abhängigkeit der situativen, technischen, wirtschaftlichen und rechtlichen Rahmenbedingungen (Hauß & Timpe, 2000). Außerdem sind die Personenmerkmale des Entwicklers bedeutsam (Hacker, 1998; Riemer, Weißhahn, & Hacker, 1999).

Hinsichtlich der Personenmerkmale kommt der Kapazität des Arbeitsgedächtnisses eine besondere Rolle zu (Kyllonen & Christal, 1990). Es stellt für die jeweilige Aufgabe relevantes Wissen aus dem Langzeitgedächtnis bereit, bearbeitet dieses sowie die neu erworbenen Kenntnisse und speichert den neuen Wissens- und Erkenntnisstand ab (Dörner, 1995; Lürer & Spada, 1990). Somit ist das Arbeitsgedächtnis eine kontrollierende, koordinierende und verarbeitende Instanz (Dörner, 1995). Bei dem relevanten Wissen handelt es sich um Sach- und Methodenwissen (Hacker & Sachse, 2006). Durch eine Überbelastung des Arbeitsgedächtnisses kann es zu Problemen bei der Bereitstellung, Bearbeitung und Speicherung von Wissen kommen, dessen Konsequenz Denkfehler (Ehrlenspiel, 1993), das Übersehen wichtiger Aspekte oder eine zu starke Abstraktion des Problems (Ehrlenspiel, 2003) sowie Fehlentscheidungen

(Weißhahn & Rönsch, 2002) sein können. Ebenso kann ein überlastetes Arbeitsgedächtnis beim Entwickler zu vermehrt unnötigen Reflexionen und Handlungen führen, wodurch dieser insgesamt mehr Zeit für die Entwicklung benötigt (Sachse & Leinert, 1996). Häufig wird zur Entlastung des Arbeitsgedächtnisses paralleles Bearbeiten einzelner Teilaufgaben vermieden und auf Altbewährtes beim Vorgehen oder als Lösung zurückgegriffen (Schütze, 2003). Wie schwierig eine Entwicklungstätigkeit ist, ergibt sich somit nicht allein aus den Merkmalen des zu entwickelnden Systems, sondern zudem aus den Leistungsfähigkeiten und -fertigkeiten des Entwicklers (Hacker, 1998; Schmalt & Langens, 2009).

#### 4.2.1.1 Vorgehensmerkmale erfolgreicher Entwickler

Ein erfolgreicher Entwickler zeichnet sich durch folgende Vorgehensmerkmale bei seiner Tätigkeit aus:

**Anforderungsanalyse:** Zu Beginn der Entwicklung wird eine detaillierte Analyse der Aufgaben, des Einsatzbereiches, der bereits existierenden ähnlichen Systeme und Anforderungen (Dylla, 1991) sowie ggf.<sup>79</sup> der Benutzer des zu entwickelnden Systems durchgeführt (Sarodnick & Brau, 2006). Dadurch lässt sich ein umfassendes Problemverständnis gewinnen (Christiaans & Dorst, 1992). Aus dieser Analyse werden zudem weitere Anforderungen an das zu entwickelnde System abgeleitet, welche wiederum die Basis für dessen Funktionalität sowie ggf.<sup>80</sup> das Anzeige- und Bedienkonzept bilden (Sarodnick & Brau, 2006). Diese Analyse wird zur Absicherung der bestehenden Anforderungen sowie zur Aufdeckung neuer Anforderungen im Laufe der Systementwicklung wiederholt (Dylla, 1991), vor allem aufgrund der gewonnenen neuen Erfahrungen und Informationen.

**Wissensansammlung:** In die Entwicklung werden vielfältige Informationen und Wissen, vor allem über die Funktionalität des Systems, einbezogen (Görner, 1994; Hacker, Wetzstein, & Römer, 2002).

**Zielgerichtetheit:** Die Entwicklertätigkeiten werden auf die Entwicklung des Systems ausgerichtet und dafür zügig ausgeführt (Dylla, 1991; G. Fricke, 1993; Hacker & Sachse, 2006).

**Flexibilität:** Das Vorgehen wird flexibel an die sich im Laufe der Entwicklungszeit ändernden Anforderungen und neu erworbenen Erkenntnisse angepasst (G. Fricke, 1993).

---

<sup>79</sup> Im Falle der Entwicklung eines vollautomatischen Systems existieren keine Benutzer, welche einer Analyse unterzogen werden können. Es sei denn, für diese Entwicklungsaufgabe kann der Benutzer eines ähnlichen bestehenden MMS relevante Informationen liefern, etwa zu seinen typischen Arbeitsschritten oder welche Informationen der Umwelt er zur Aufgabenausführung heranzieht.

<sup>80</sup> Ein Anzeige- und Bedienkonzept ist im Falle der Entwicklung eines vollautomatischen Systems nicht relevant. Es sei denn, es wird bspw. für Wartungs-/Reparaturarbeiten eine Mensch-Maschine-Schnittstelle mit Diagnoseinformationen für den Techniker sowie für Korrekturangaben durch den Techniker vorgesehen.

**Wechsel zwischen Abstraktionsebenen:** Zunächst erfolgt eine grobe Bearbeitung des Gesamtsystems (Hacker et al., 2002). Dem schließt sich eine detaillierte Entwicklung einzelner Teilaspekte an (Hacker et al., 2002). Im Verlauf der Entwicklungstätigkeit findet immer wieder ein Wechsel zwischen der Bearbeitung des Gesamtsystems und den Teilaspekten statt (G. Fricke, 1993), bis schließlich das Gesamtsystem fein ausgearbeitet ist.

**Mehrere Lösungsoptionen:** Auf den verschiedenen Abstraktionsebenen werden mehrere Lösungsalternativen entwickelt (Hacker et al., 2002).

**Erfolgskontrollen:** Die so generierten Lösungsoptionen werden hinsichtlich der Anforderungserfüllung überprüft (Kelsch, Flemisch, Löper, Schieben, & Schindler, 2006), die geeignetste(n) Variante(n) ausgewählt (Ehrlenspiel, 2003) und weiter optimiert (Sarodnick & Brau, 2006). Diese beständigen analytischen Erfolgskontrollen sowie die kritischen Betrachtungen der eigenen Vorgehensweisen sind entscheidend für die weiteren Handlungen (Eisentraut & Günther, 1997). Erst wenn eine Idee bzw. Option der Evaluation standhält, wird sie umgesetzt, optimiert und erneut getestet (Badke-Schaub, 2005). Um Zeit zu sparen, wird jedoch oftmals in der Produktentwicklung auf die analytische Erfolgskontrolle verzichtet, wodurch meist vorschnell ungeeignete Lösungen ausgebaut oder gute Lösungen abgelehnt werden (Badke-Schaub & Frankenberger, 2004).

**Reflexiver äußerer Dialog:** Eine wichtige Methode zur Generierung und Überprüfung der einzelnen Lösungsoptionen ist das „sprechende Entwickeln“ in Form von Gesprächen und Diskussionen mit Kollegen, Auftraggebern, aber auch mit unbeteiligten, unwissenden und Fragen stellenden Personen (Hacker & Sachse, 2006). Dabei ist der Dialog in Form von Erläuterungen, Bewertungen und Schlussfolgerungen entscheidend für eine Verbesserung der Güte des Systems (Görner, 1994; Hacker & Wetzstein, 2004). Einen Beitrag zur positiven Wirkung des reflexiven Dialogs können die Frage-Antwort-Techniken leisten (Winkelmann & Hacker, 2006), im Speziellen die W-Fragen (Hacker & Sachse, 2006).

**Reflexiver innerer Dialog:** Das „sprechende Entwickeln“ in Form eines inneren Bewertungsmonologs spielt ebenso eine wichtige Rolle für die Güte des Systems (Hacker & Wetzstein, 2004). So konnten Hacker et al. (2002) aufzeigen, dass ein höherer Zeitanteil ohne ersichtliche Aktivität der Entwickler während der Entwicklungstätigkeit mit einer höheren Güte des entwickelten Systems einhergeht. Die Autoren interpretierten diesen Zeitanteil als ein Nachdenken mit reflexivem Inhalt über die eigene Arbeit.

**Externalisierung:** Durch externe abstrakte Darstellungen von Gedächtnisinhalten können das Arbeitsgedächtnis entlastet, Denkprozesse unterstützt (Ehrlenspiel, Bernard, & Günther, 1996; Hacker, 1998) und die Komplexität des zu bearbeitenden Problems reduziert

werden (Hacker & Sachse, 2006). Dadurch können neue Lösungsoptionen angeregt, entwickelt, weiter bearbeitet, konkretisiert, bewertet und vor dem Vergessen geschützt werden (Hacker & Sachse, 2006). Für die externe Fixierung eignen sich das schriftliche Entwickeln von Lösungsoptionen, das schriftliche Ausfüllen von Checklisten sowie das Skizzieren (Hacker & Sachse, 2006).

#### 4.2.2 Merkmale der Entwicklungsaufgabe

Mängel bei der Entwicklung ergeben unzuverlässige technische Systeme (Sicinski, 1995). Das ist trotz der Anwendung geeigneter Vorgehensweisen nicht immer zu vermeiden, denn eine Entwicklungsaufgabe ist, aufgrund ihrer nachfolgend dargestellten Merkmale, anspruchsvoll:

**Ungenaue/konträre Anforderungen:** Häufig liegen nur unvollständige, ungenaue (G. Fricke, 1993; Winkelmann & Hacker, 2006) und teilweise konträre (Badke-Schaub & Frankenberger, 2004; Visser, 1994) Anforderungen an das zu entwickelnde System vor. Bei den Anforderungen handelt es sich um objektive Eigenschaften, welche die Entwicklungstätigkeit und das -ergebnis erfüllen müssen (Hacker & Sachse, 2006).

**Ungenauer/widersprüchlicher Zielzustand:** Aus diesen unpräzisen Anforderungen kann oftmals der Zielzustand des technischen Systems ebenfalls nur als ungenaue oder sogar widersprüchliche Beschreibung abgeleitet werden (Hacker & Sachse, 2006; Pietzcker, 2004; Winkelmann & Hacker, 2006).

**Hohe Interdependenz:** Ebenso ergeben sich aus den Anforderungen die zu bewältigenden zumeist vielfältigen Teilaufgaben während der Entwicklungstätigkeit (Dylla, 1991; G. Fricke, 1993). Diese sind häufig interdependent (Sarter et al., 1997), d. h. deren Bearbeitung beeinflusst wiederum die Lösung einer anderen Teilaufgabe. Oft müssen einzelne Teilaufgaben daher parallel bewältigt werden (G. F. Smith & Browne, 1993).

**Hohe Komplexität:** Die Anzahl der Teilaufgaben, deren Verknüpfungsintensität und gleichzeitige Bearbeitung bestimmen die Komplexität einer Entwicklungstätigkeit (Dörner, 1987; Hacker & Sachse, 2006). Ein zu entwickelndes technisches System ist zumeist komplex (Hacker & Sachse, 2006; Rocholl, 2010; Sarter et al., 1997), wobei dessen Komplexität durch die wachsenden technischen Möglichkeiten stetig zunimmt.

**Anforderungsveränderungen:** Die Anforderungen an ein System können im Laufe der Entwicklung immer wieder Änderungen unterliegen oder ausdetailliert werden (Badke-Schaub & Frankenberger, 2004; Rupp, 2006; Sarodnick & Brau, 2006; VDI, 1993), wodurch sich Entwickler mehrfach mit einzelnen Aspekten und deren Lösungsoptionen beschäftigen müssen. Eine vollständige und konkrete Spezifikation des zu entwickelnden Systems, mit all seinen

einzelnen Funktionen, Wirkmechanismen und Parametern, kann somit erst nach und nach im Entwicklungsprozess erstellt werden. Ebenso liegen den Entwicklern zum Entwicklungsbeginn häufig unvollständige Informationen über die situativen, finanziellen sowie rechtlichen Rahmen- und Umweltbedingungen sowie den Vorgehensmöglichkeiten vor. Dabei werden Entwickler oft durch zeitliche Veränderungen der Rahmen- und Umweltbedingungen mit einer Eigendynamik der Entwicklungsaufgabe konfrontiert (Hacker & Sachse, 2006).

**Interdisziplinarität:** Die Entwicklung von Systemen findet immer häufiger, zumindest phasenweise, in interdisziplinären Teams statt, um die verschiedenen und umfangreichen Erfahrungs- sowie Wissenshintergründe (Dörner, 1989; Ehrlenspiel, 2003) und Kompetenzen der Teammitglieder zu nutzen (Hacker & Sachse, 2006; Herczeg, 2009). Durch die sich daraus ergebenden Synergieeffekte (Scherer, 1998) kann die hohe Komplexität dieser Systeme und deren Funktionalität bewältigt werden (Burns & Vicente, 1995). Erst durch die interdisziplinäre Arbeitsteilung ist es zumeist möglich, die vom Auftraggeber oder Markt angeforderte Systemqualität zu erreichen und den angezielten Kosten- und Zeitaufwand einzuhalten (Frankenberger, 1997). Dafür ist eine Wissensintegration und die Ausbildung einer gemeinsamen Wissensbasis sowie eines Problem-, Ziel- und Vorgehensverständnisses von Nöten (Anderson, 2001; Badke-Schaub, 2005; Dorst, 2003).

Für die Entwicklung von komplexen, technischen Systemen reicht ein Team zumeist nicht aus. Verschiedene Teams mit unterschiedlicher Expertise entwickeln und bearbeiten Teilaspekte des Systems. Diese Teams können verschiedenen Abteilungen oder gar Unternehmen angehören. Dabei steigt der Koordinationsaufwand umso mehr an, je mehr arbeitsteilige Personen, Teams, Abteilungen oder Unternehmen in den Entwicklungsprozess involviert sind. Diese Einheiten müssen im ständigen Kontakt zum Austausch und zur Abstimmung stehen. Hierbei sind Kompromissbereitschaft und größtmögliche Offenheit für neue Vorgehensweisen notwendig, um gemeinsam ein optimales System entwickeln zu können. Eine ungenaue Rollendefinition und eine konkurrenzfördernde Struktur sind zu vermeiden, da sie die Bereitschaft zur Kooperation und zum Austausch verringern (Badke-Schaub & Frankenberger, 2004).

Trotz interdisziplinärer Entwicklungsteams werden einzelne Arbeitsschritte im Entwicklungsprozess zum Nutzen spezieller Fachkenntnisse nur durch einen Entwickler allein ausgeführt (Badke-Schaub & Frankenberger, 2004; G. Fricke, 1993; Günther, 1998; Hacker, 1998). Das dient der detaillierten Ausarbeitung, Implementierung und/oder Prüfung von Lösungsansätzen. Diese „Solotätigkeit“ nimmt einen beachtlichen Anteil der gesamten Entwicklungszeit ein und spart dennoch oft Zeit ein (Badke-Schaub & Frankenberger, 2004).

Bei der Entwicklung von MMS oder deren Teilkomponenten, den Assistenzsystemen oder Automaten, handelt es sich somit um einen iterativen (Hacker & Sachse, 2006), dynamischen, komplexen, innovativen (Hacker et al., 2002; Rocholl, 2010), kreativen (Dörner, 1989), risikohaften und unsicherheitsbelasteten Prozess (Emmanuelides, 1993).

#### 4.2.2.1 Design Problem Solving

Aufgrund der Merkmale einer Entwicklungsaufgabe wird das Entwickeln von komplexen, dynamischen, technischen Systemen oftmals mit dem komplexen Problemlösen (engl. Design Problem Solving) gleichgesetzt (Badke-Schaub & Frankenberger, 2004; Cross, Christiaans, & Dorst, 1996; G. Fricke, 1993; Hacker & Sachse, 2006; Hacker et al., 2002; Johannsen, 1993; Pietzcker, 2004; Simon, 1995; G. F. Smith & Browne, 1993; P. Smith, McCoy, & Layton, 1997; Winkelmann & Hacker, 2006).

Wie der Tabelle 20 zu entnehmen ist, sind die Merkmale eines komplexen Problems, aufgestellt durch Dörner (1989), vergleichbar mit denen einer Entwicklungsaufgabe.

*Tabelle 20: Merkmale von komplexen Problemen nach Dörner (1989), deren Bedeutungen sowie äquivalente Merkmale von Entwicklungsaufgaben*

Merkmalsbedeutung	Merkmale eines komplexen Problems nach Dörner (1989)	äquivalentes Merkmal der Entwicklungsaufgabe
Nicht alle lösungsrelevanten Informationen sind dem Problemlöser bekannt.	<b>Intransparenz</b>	ungenauere Anforderungen
Es liegt ein ungenauer Zielzustand vor, der durch den Problemlöser konkretisiert werden muss.	<b>Offenheit</b>	ungenauer Zielzustand
Es müssen mehrere, teilweise kontradiktorische Ziele erreicht werden.	<b>Polytelie</b>	vielfältige, teilweise konträre Teilaufgaben
Die einzelnen Variablen sind stark miteinander vernetzt, wobei dies dem Problemlöser nicht immer offensichtlich ist. Die Beeinflussung einer Variable führt zu Veränderungen anderer Variablen.	<b>Vernetztheit</b>	Interdependenz
Es müssen immer mehrere Variablen berücksichtigt werden. Die Anzahl der vorhandenen Variablen bestimmt die Komplexität.	<b>Komplexität</b>	Komplexität
Zustandsänderungen des Systems geschehen ohne Beeinflussung durch den Problemlöser.	<b>Eigendynamik</b>	Anforderungsveränderung

Nach Dörner (1987) existiert bei einem Problem ein unerwünschter Ausgangs- und ein erwünschter Zielzustand. Auf dem Weg vom Ausgangs- zum Zielzustand liegen mittels Lösungsoperatoren zu überwindende Barrieren. Der Ausgangs-, der Zielzustand, die Lösungsmittel und deren vielfältige Kombinationen stellen den Problemraum dar, innerhalb dessen eine adäquate Lösung gefunden werden muss. Dabei wird in Abhängigkeit davon, welche Barriere zur Lösungserreichung bewältigt werden muss, zwischen vier Problemtypen unterschieden:



**1. Problem mit Interpolationsbarriere:** Die Ziele sind dabei eindeutig und die notwendigen Operatoren zur Zielerreichung bekannt. Jedoch müssen deren Anwendungsreihenfolge und Kombination vom Problemlöser selbst identifiziert werden.

**2. Problem mit Synthesebarriere:** Dieses besteht, wenn die Zielkriterien klar formuliert, jedoch die Mittel zur Zielerreichung kaum bekannt oder nicht vorhanden sind. Sie müssen erst vom Problemlöser synthetisiert werden.

**3. Problem mit dialektischer Barriere:** Hierbei dagegen sind zwar die Methoden zur Zielerreichung bekannt, das Ziel selbst aber nur durch vage Kriterien beschrieben oder es herrscht unzureichendes Wissen darüber vor. Der Problemlöser muss den Zielzustand erst in einem dialektischen Prozess erarbeiten.

**4. Problem mit dialektischer und Synthesebarriere:** Bei diesem sind sowohl die Zielkriterien unpräzise als auch die notwendigen Operatoren unbekannt oder nicht vorhanden. Der Problemlöser muss den Zielzustand und die dahin führenden Operatoren ermitteln.

Bei der Entwicklung von technischen Systemen, wie Automaten oder Assistenzsystemen, liegen häufig nur unzureichende Informationen über den Zielzustand vor, aber auch die Methoden zur Zielerreichung sind größtenteils unbekannt oder nicht vorhanden (G. Fricke, 1993; Pietzcker, 2004). Folglich muss hierfür von den Entwicklern zumeist der schwerste Problemtyp mit der dialektischen und Synthesebarriere überwunden werden, um einen unerwünschten Ausgangszustand in einen erwünschten Zielzustand zu überführen (Rocholl, 2010). Bei dieser Tätigkeit begegnen Entwickler, aufgrund der unterschiedlichsten Anforderungen und den sich daraus ergebenden verschiedensten Teilzielen eines Systems, ebenfalls den anderen drei Problemtypen respektive wandeln einige dialektische und Synthesebarrieren im Laufe des Entwicklungsprozesses erst in einfachere Problemtypen um (G. Fricke, 1993), bis letztlich alle Barrieren überwunden sind und ein fertiges System entstanden ist. Ebenso kann erst im Laufe der Entwicklungstätigkeit durch sich ändernde Anforderungen ein Problem mit der dialektischen und der Synthesebarriere entstehen.

Das Vorgehen zur Lösung eines komplexen Problems, somit auch einer Entwicklungsaufgabe, kann über den iterativen Problemlösezyklus mit sechs Schritten vom Verein Deutscher Ingenieure (1993) beschrieben werden (Dylla, 1991; G. Fricke, 1993):

**1. Problemanalyse:** Zuerst werden die notwendigen Informationen zum Verständnis des Problems zusammengetragen und daraus die Anforderungen identifiziert.

**2. Problemformulierung:** Anschließend wird eine Konkretisierung des Problems vorgenommen.

**3. Systemsynthese:** Daraufhin werden verschiedene Ideen generiert sowie bestehende Lösungen herangezogen und miteinander verknüpft.

**4. Systemanalyse:** In diesem Schritt werden die Lösungsideen und deren Funktionsmerkmale kritisch auf ihre Tauglichkeit durchleuchtet.

**5. Beurteilung:** Dem folgt die Bewertung der Lösungsideen als Ergebnis der vorhergehenden Analyse.

**6. Entscheidung:** Abschließend findet die Auswahl der geeignetsten Lösung für das bestehende Problem statt.

#### 4.2.3 Entwicklungsprozess

Neben diesem Problemlösezyklus (VDI, 1993) existieren verschiedene weitere Modelle zur Beschreibung des systematischen Vorgehens bei der Entwicklung von technischen Systemen. Bei diesen handelt es sich um vereinfachte, präskriptive Modelle, welche dem Entwickler einen strategischen Handlungsablauf darreichen, um den Entwicklungsprozess zu optimieren (G. Fricke, 1993) und zu strukturieren. Dadurch ist die Aufgabenkomplexität leichter zu beherrschen.

Die Modelle sind im Rahmen verschiedener Disziplinen für unterschiedliche Produkte entstanden. So sind bekannte Vorgehensmodelle in der Informatik für die Softwareentwicklung das Wasserfall-Modell (Royce, 1970), das V-Modell (BIT, 2015), das Spiralmodell (Boehm, 1988), das evolutionäre/inkrementelle Modell (Sommerville, 2007), das Prototypen-Modell (Balzert, 2000) und das Concurrent Engineering Model (Endsley, Bolté, & Jones, 2003). In den Ingenieurwissenschaften sind für die Produktentwicklung bspw. das Modell des Vereins Deutscher Ingenieure (1993), das Modell der Produktentwicklung von Pahl und Beitz (Pahl, Beitz, Feldhusen, & Grote, 2007) und das Modell des konstruktiven Entwurfsprozesses von Ehrlenspiel (2003) verbreitet. Speziell zur Bestimmung des angemessenen Ausmaßes der Automation für ein spezifisches MMS bieten Parasuraman et al. (2000) ein Vorgehensmodell im Rahmen eines iterativen Designprozesses an. Für die konkrete Entwicklung eines gebrauchstauglichen Anzeige- und Bedienkonzepts in einem MMS sind im Rahmen des Usability-Engineering anerkannte Modelle entstanden, etwa das Usability-Engineering Lifecycle Model (Mayhew, 1999) und das Prozessmodell Usability-Engineering (Sarodnick & Brau, 2006). Das Usability-Engineering findet jedoch nicht immer ausreichend Beachtung bei der Entwicklung von interaktiven Systemen (Sarodnick & Brau, 2006). Da sich das Usability- und Product-Engineering gegenseitig bedingen, sollten diese parallel, in enger Verknüpfung durchgeführt werden (Sarodnick & Brau, 2006). Um sowohl die technischen Komponenten als auch die Mensch-Maschine-Interaktion von Anfang an in einem Systementwicklungsprozess

zu berücksichtigen, ist aus der Ingenieurspsychologie das Modell der parallel-iterativen Systemgestaltung hervorgegangen (Timpe & Kolrep, 2000).

Den Modellen gemein ist, dass sie die Phasen von der „Problem- und Anforderungsidentifizierung“ bis hin zum „marktreifen System“ umfassen. Einige von ihnen beginnen noch früher oder enden später im Produktkreislauf, so z. B. vom Bedarf bis zur Verschrottung respektive zum Recycling. Die einzelnen Modelle unterscheiden sich zudem in ihrer Komplexität, der Anzahl der definierten Entwicklungsphasen sowie in ihrer konkreten Bedeutung und Bezeichnung. Aus diesen Modellen abgeleitet, lassen sich zur Vereinfachung alle Arbeitsschritte für die Entwicklung eines technischen Systems in vier Phasen einteilen: die „Analysephase“, die „Konzeptphase“, die „Entwicklungsphase“ und die „Einführungsphase“.<sup>81</sup>

**Analysephase:** Diese Phase dient der Sammlung und umfassenden Sichtung aller relevanten Informationen und ähnlicher Systeme zur Präzisierung der Aufgabenstellung sowie Identifizierung der Anforderungen an das System und die Entwicklungsbedingungen.

**Konzeptphase:** In dieser Phase werden die prinzipiellen Lösungen in Form von Funktionsstrukturen und Wirkmechanismen generiert. Dafür wird das System in Teilaspekte aufgeschlüsselt. Für die Teilaspekte werden verschiedene Lösungsoptionen aufgestellt. Diese werden bis hin zum Gesamtsystem miteinander kombiniert, kritisch auf die Erfüllung der Anforderungen bewertet und schließlich die geeignetsten für die Entwicklung ausgewählt.

**Entwicklungsphase:** In der Entwicklungsphase findet die Verfeinerung des Lösungskonzepts statt. Hierbei wird eine konkrete, realisierbare Umsetzung des Gesamtsystems unter Berücksichtigung technischer und wirtschaftlicher Aspekte erarbeitet und bewertet. Dabei wird die Funktionalität, Zuverlässigkeit, Gebrauchstauglichkeit sowie Beherrschbarkeit des Gesamtsystems ermittelt. Dem folgt die Systemdefinition und -freigabe für die Serienproduktion.

**Einführungsphase:** In der Einführungsphase wird das System produziert und in den Markt eingeführt.

In der Literatur lassen sich ähnlich stark vereinfachte Darstellungen des Entwicklungsprozesses mit drei bis fünf Phasen finden (s. Ehrlenspiel, 2003; Frankenberger, 1997; Pahl et al., 2007; Reichart, 2002; Sarodnick & Brau, 2006; Schütze, 2003). Eine noch stärkere Vereinfachung unterteilt den Prozess in zwei Hauptphasen: die „Konzeptionsphase“ und die „Entwicklungsphase“ (s. Peeters, Rutte, Tuijl, & Reymen, 2008; J. Schwarz, 2007). Wobei sich die Konzeptionsphase aus der Analyse- und der Konzeptphase zusammensetzt. Sie entspricht somit

---

<sup>81</sup> Entspricht den Phasenbezeichnungen des Usability-Engineering-Modells (Sarodnick & Brau, 2006), welches sich jedoch nur auf die Entwicklung einer gebrauchstauglichen Mensch-Maschine-Schnittstelle bezieht.

den frühen Phasen der Systementwicklung (Verworn & Herstatt, 2007). In ihr findet vorrangig divergentes Denken statt (Eris, 2003), welches im Zusammenhang mit Offenheit für Erfahrung und Kreativität steht (McCrae, 1987; Runco & Okuda, 1988). Diese Denkweise dient der Generierung von vielen Lösungsoptionen (Baya & Leifer, 1996). Dagegen dominiert in der Entwicklungsphase das konvergente Denken (Eris, 2003). Dabei handelt es sich um analytisches, rationales und logisches Denken, welches für die Bewertung (Cropley, 2006) und ggf. Synthese (Ehrlenspiel, Kiewert, & Lindemann, 2005) der Lösungskonzepte ausgeführt wird. Beide Denkweisen sind für die Entwicklung eines erfolgreichen, neuartigen Systems notwendig (Cropley, 2006). Dagegen führt ein Verharren in divergenter Denkweise i. d. R. zu schlechteren Systementwicklungen (Eris, 2003).

Im Sinne der Bearbeitung aller Problemaspekte auf einem Abstraktionsniveau vor einem Wechsel zu einer höheren Granularität konnte Fricke (1993) empirisch das phasenweise Vorgehen bei einer Entwicklungstätigkeit nachweisen. Dabei zeigte sich auch, dass ein flexibler Umgang mit dem durch ein Vorgehensmodell standardisierten Entwicklungsablauf gewinnbringender ist als dessen starre Umsetzung. Die Übergänge zwischen den einzelnen Phasen des Produktkreislaufs sind fließend. Zudem werden die Phasen zumeist nicht rein sequentiell durchlaufen, sondern unterliegen einem iterativen Prozess und finden teilweise parallel statt (Brodbeck, 2001; Mayhew, 1999; Preim, 1999; Reichart, 2002; Sarodnick & Brau, 2006; VDI, 1993). Dabei ergibt sich das iterative Vorgehen durch die Entwicklung-Bewertung-Fehlerkorrektur-Schleifen (Sarodnick & Brau, 2006; G. F. Smith & Browne, 1993). Dieses flexible Vorgehen wird in den oben genannten Modellen in unterschiedlichem Ausmaß berücksichtigt. In der Realität herrschen jedoch oftmals recht unflexible Prozesse bei der Entwicklung vor (Pahl et al., 2007), ohne oder mit unzureichenden Rückkopplungsschleifen. Zudem besteht in der Theorie und der Praxis Uneinigkeit darüber, ob einzelne Phasen ganz übersprungen werden dürfen. So rät der Verein Deutscher Ingenieure in der Richtlinie 2221 (1993) davon ab, eine Phase zu übergehen. Dagegen zeigen Untersuchungen, dass das Überspringen einzelner Phasen teilweise sogar zu besseren Ergebnissen führt (Visser, 1994) und kein untypisches Verhalten im Entwicklungsprozess ist (Badke-Schaub & Stempfle, 2003).

Prinzipiell ist ein flexibles Vorgehen im Entwicklungsprozess von Nöten, um neu gewonnenem Wissen und gesammelten Erfahrungen (Simon, 1995), neuen (Badke-Schaub & Stempfle, 2003) sowie sich geänderten Anforderungen gerecht werden zu können.

#### **4.2.3.1 Systematisches vs. opportunistisches Vorgehen bei der Entwicklung**

Des Weiteren kann eine Entwicklungstätigkeit durch opportunistisches Verhalten der Entwickler geprägt sein (Günther & Ehrlenspiel, 1998; Hacker, 1996; Hacker et al., 2002).

Bevor zur nächsten Ebene gewechselt wird, findet hierbei keine systematische Zerlegung der Entwicklungsaufgabe in Unteraufgaben und die Bearbeitung aller Aspekte auf einer Abstraktionsebene zur Identifizierung von Lösungsoptionen statt (Hacker & Sachse, 2006; Hacker et al., 2002). Vielmehr verwenden Entwickler ihnen bekannte, nutzbare Lösungen, welche sie bereits im Gedächtnis abgespeichert haben (Hacker & Sachse, 2006). Findet sich eine solche Lösungsoption für einen Teilaspekt oder das gesamte System, wird diese ausgearbeitet und umgesetzt, ohne eine weitere Analyse des Gesamtproblems oder die Entwicklung weiterer Lösungsalternativen in Betracht zu ziehen (Visser, 1994; von der Weth, 2001). Dabei richtet sich die Reihenfolge der zu bearbeitenden Systemaspekte nach dem vorhandenen Wissen und der Erfahrung der Entwickler (Hacker & Sachse, 2006). Nacheinander bearbeitete Teilaspekte müssen demzufolge in keinem direkten Zusammenhang stehen (Hacker & Sachse, 2006). Ebenso können Sprünge zwischen verschiedenen Teilaspekten sowie dem Gesamtsystem, aber auch den verschiedenen Lösungsalternativen stattfinden (Hacker et al., 2002). Durch dieses erfahrungsgeleitete Handeln bei der Entwicklung lassen sich unterschiedlichste Vorgehensweisen finden (Hacker & Sachse, 2006; Hacker et al., 2002). Weder eine rein opportunistische, noch eine systematische Vorgehensweise weisen eine Allgemeingültigkeit auf (Hacker & Sachse, 2006; Hacker et al., 2002; von der Weth, 2001). Das obgleich die Merkmale des erfolgreichen Entwickelns, wie das Generieren mehrerer Lösungsalternativen, am Stärksten bei der systematischen Vorgehensweise vertreten sind. Ebenso zeigen sich Entwickler durch eine solche Vorgehensweise am zufriedensten. Gleichzeitig fühlen sich diese in ihrem Vorgehen stark beschränkt und belastet (Hacker et al., 2002).

Oft sind bei einer Entwicklungstätigkeit verschiedene Vorgehensweisen und Lösungen möglich (Hacker, 1998). Als besonders geeignet erweisen sich Mischformen aus systematischem und opportunistischem Vorgehen (Hacker, 1989). So fanden Smith und Brown (1993) einen Vorteil für die systematische Zerlegung des Gesamtproblems in Teilaspekte bis diese durch Wissen gelöst werden können. Ebenso konnte Visser (1994) den positiven Einfluss einer opportunistischen Strategie mit einzelnen systematischen Phasen auf die Leistungsgüte aufzeigen. Erfahrene Entwickler zeigen häufig einen Wechsel zwischen einer systematischen Aufteilung der Gesamtaufgabe in Teilprobleme und einem opportunistischen Vorgehen zur Lösung einzelner Teilaspekte (Hacker & Sachse, 2006). Diese Mischstrategien bieten eine adäquate Flexibilität im Entwicklungsprozess (Hacker et al., 2002).

#### **4.2.3.2 Frühe Phasen des Entwicklungsprozesses**

Den frühen Phasen des Entwicklungsprozesses, bis zum Ende der Konzeptionsphase, wird eine besondere Rolle für den Entwicklungserfolg zugeschrieben. In diesen Phasen ist der

Einfluss auf ein flexibles Vorgehen, die Fehleranzahl, die Qualität, die Innovation und die Kosten des Systems am höchsten (Auer & von der Weth, 1994; Baya & Leifer, 1996; Dylla, 1991; Ehrlenspiel, 2003; Ehrlenspiel & Günther, 1995; Ehrlenspiel et al., 2005; G. Fricke, 1993, 1994; Hacker, 1994; Hacker & Sachse, 2006; Rocholl, 2010). Gleichzeitig lassen sich in dieser Entwicklungsphase die praktische Umsetzbarkeit (Hacker & Sachse, 2006; Rocholl, 2010) und die Kosten für das System noch schlecht einschätzen (Ehrlenspiel, 2003). Sichere Schätzungen können meist erst in späteren Entwicklungsphasen getroffen werden. Dann gibt es jedoch i. d. R. kaum mehr Möglichkeiten, einen geeigneteren respektive günstigeren Weg einzuschlagen. Deshalb ist es wichtig, bereits in den frühen Phasen der Entwicklung diese Unsicherheiten weitestgehend zu reduzieren (Verworn, 2005).

Zudem verursachen in diesem Entwicklungszeitraum Fehlerkorrekturen wenig Aufwand und die geringsten Kosten. Im weiteren Entwicklungsprozess steigen der Aufwand und die Kosten zur Fehlerbehebung kontinuierlich an (Ehrlenspiel, 2003; Ehrlenspiel et al., 2005).

## 5. Determinanten der Leistung von Entwicklern

Dieses Kapitel widmet sich einem Teilziel des Projekts ATEO, der Identifizierung der Leistungs determinanten von Entwicklern von komplexen, dynamischen, technischen Systemen. Die Leistungs determinanten werden nach der Arbeitsdefinition des Projekts als Faktoren verstanden, die jeweils einen positiven oder aber negativen Einfluss auf die Leistung der Entwickler haben. Dabei handelt es sich um Faktoren, welche den einzelnen Entwicklern zuzuschreiben sind (etwa Persönlichkeitsmerkmale oder die Expertise), welche aber auch von außen auf sie einwirken (bspw. die Teamgröße oder der Zeitdruck). Wenn die jeweilige optimale Ausprägung dieser Leistungs determinanten bekannt ist, kann die Arbeit von Entwicklern zukünftig besser unterstützt und deren Leistung optimiert werden.

Nachfolgend soll, zum besseren Verständnis der Vorgehensweise bei der Identifizierung der Leistungs determinanten, ein kurzer Rückblick über die dazugehörigen Forschungsbemühungen in der ersten ATEO-Projektphase gegeben werden. Dem schließt sich eine ausführliche Beschreibung des Vorgehens der Autorin sowie der dadurch gewonnenen Ergebnisse an.

### 5.1 ATEO-Projektphase I: Identifizierung der Leistungs determinanten

In der ersten ATEO-Projektphase fanden zwei experimentelle Untersuchungen auf Seiten des Entwicklers zur systematischen Ermittlung seiner Leistungs determinanten statt (Krinner, 2009). Diese beiden Untersuchungen werden im Folgenden kurz beschrieben.

#### 5.1.1 Untersuchung Leistungs determinante „Art der Information über das System“

Die erste Untersuchung diente der Ermittlung des Einflusses der uV „Art der Information über das System“ auf den Entwicklungsprozess und das -ergebnis (Krinner, 2007a, 2007b). Es nahmen 31 Teams à zwei Personen teil. Bei den Teilnehmern handelte es sich um Studierende technischer und ingenieurwissenschaftlicher Studiengänge, die sich bereits im Hauptstudium befanden. Die Ausprägung der uV wurde wie folgt in einem Between-Groups-Design dreifach variiert:

1. *Bedingung: vereinfachtes Lastenheft (Beschreibung kooperative Tracking-Aufgabe und zu bewältigendes Problem; zehn Teams),*
2. *Bedingung: vereinfachtes Lastenheft + eigene Durchführung der kooperativen Tracking-Aufgabe (zehn Teams),*
3. *Bedingung: vereinfachtes Lastenheft + Betrachten der MWB beim Ausführen der kooperativen Tracking-Aufgabe per Screenrecording-Video + MWB-Interview (elf Teams).*

Für die Überprüfung, welche Art von Assistenzsystemen konzipiert wurden, sind diese in die Taxonomie für Assistenzsysteme von Wandke (2005) und Wetzenstein (2004, s. Abschnitt 4.1.5) eingeordnet worden. Es zeigten sich keine signifikanten Unterschiede zwischen den drei Ausprägungsstufen der uV „Art der Information über das System“ bzgl. der Anzahl der Assistenzsystemkonzepte, der Anzahl der unterstützten Handlungsphasen, der Anzahl an Assistenzsystemen je Handlungsphase und den weiteren Dimensionen der verwendeten Taxonomie (mit Ausnahme beim angedachten Medium für eine Entscheidungsassistenz) sowie der Art der definierten Aufgabenteilung für die MWB je Entwicklerteam. Für die Art der Unterstützung von Entscheidungen der MWB konnte jedoch ein signifikanter Unterschied aufgedeckt werden. Von den Teams der Bedingung 2, welche selbst kooperativ das Tracking ausführen durften, wurden weniger explizite Regelungen konzipiert als von den Teams der anderen Bedingungen.

Krinner (2009) vermutete als Ursache für die gefundenen Ergebnisse u. a. zu geringe Unterschiede zwischen den verwendeten Stufen der uV „Art der Information über das System“.

### 5.1.2 Untersuchung „Ressourcenmaximierung“

Aufgrund der kaum vorhandenen signifikanten Ergebnisse der ersten experimentellen Untersuchung wurden in einer zweiten Untersuchung von Krinner (2009) und Henkel (2007) vier uVn gleichzeitig in einem Extremgruppenvergleich hinsichtlich ihres Unterstützungspotentials beim Entwickeln von Assistenzsystemen untersucht. Jede dieser uVn wurde dabei zweifach gestuft variiert.

#### 1. Bedingung (zehn Teams):

- *Teamgröße: zwei Studenten,*
- *fachliche Teamzusammensetzung: nur Studenten einer technischen Fachrichtung,*
- *Art der Information über das System: vereinfachtes Lastenheft und*
- *Erfahrung der Teams: nur Hauptstudium.*

#### 2. Bedingung (sieben Teams):

- *Teamgröße: drei Studenten,*
- *fachliche Teamzusammensetzung: je ein Student der Psychologie + Design + technische Fachrichtung,*
- *Art der Information über das System: vereinfachtes Lastenheft + Betrachten der MWB beim Ausführen der kooperativen Tracking-Aufgabe per Screenrecording-Video + eigene Durchführung der kooperativen Tracking-Aufgabe + MWB-Interview und*
- *Erfahrung der Teams: Hauptstudium + Erfahrung und Interesse im Bereich Mensch-Maschine-Interaktion.*



Die ressourcenarme Bedingung 1 entspricht der Bedingung 1 der Untersuchung „Art der Information über das System“ von Krinner (2009). Daher wurden für diese Bedingung keine neuen Daten erhoben, sondern jene aus der vorhergehenden Untersuchung verwendet. Lediglich für die ressourcenreiche Bedingung 2 wurden in dieser zweiten Untersuchung gesondert Daten erfasst. Zur Ermittlung der Art der konzipierten Assistenzsysteme wurde wiederum die Taxonomie für Assistenzsysteme von Wandke (2005) und Wetzenstein (2004, s. Abschnitt 4.1.5) verwendet. Es zeigten sich erneut keine signifikanten Unterschiede zwischen den beiden Bedingungen der experimentellen Untersuchung „Ressourcenmaximierung“ bzgl. der Anzahl der Assistenzsystemkonzepte, der Anzahl an Assistenzsystemen je Handlungsphase und den weiteren Dimensionen der verwendeten Taxonomie sowie der Art der definierten Aufgabenteilung für die MWB je Entwicklerteam. Des Weiteren konnte diesmal bei der Art der Unterstützung von Entscheidungen der beiden MWB kein signifikanter Unterschied zwischen den beiden Gruppen ermittelt werden. Allerdings entwickelten die Teams der Experimentalgruppe im Durchschnitt im Vergleich zu denen der Kontrollgruppe Assistenzsysteme, welche eine Handlungsphase mehr unterstützen. Dabei handelte es sich vorrangig um die Effektkontrolle.

Krinner (2009) vermutete in der Interdisziplinarität den Hauptverursacher für die breitere Betrachtung des Problems und somit auch der Konzeption von umfangreicheren Unterstützungen in der ressourcenreichen Bedingung. Gleichzeitig stellte sie fest, dass aufgrund der Konfundierung der uVn die Unterschiede zwischen den beiden Gruppen nicht sicher auf eine einzelne uV zurückgeführt werden können.

### **5.1.3 Expertenevaluation der Assistenzsystemkonzepte**

Um die Güte von 30 ausgewählten Assistenzsystemkonzepten aus ihren Untersuchungen zu ermitteln, wählte Krinner (2009) die Methode der Expertenevaluation. Als Evaluatoren dienten 20 Experten aus dem deutschsprachigen Bereich der Mensch-Maschine-Interaktion mit Erfahrung bei der Gestaltung und/oder Bewertung von Assistenzsystemen. Jedes Konzept wurde durch den, von Krinner (2009) entwickelten, „Fragebogen zur Bewertung von Assistenzkonzepten“ sowie einem vergleichenden Gesamturteil, in Form einer Rangreihe der Konzepte mit Begründung für den ersten und letzten Platz, bewertet (s. Anhang F.2.1). Die anhand des Intraklassen-Korrelationskoeffizienten (ICC) berechnete Expertenkonkordanz erwies sich über alle Items des Fragebogens ( $M_{ICC} = .01$ ,  $SD_{ICC} = .23$ ) und für das Gesamturteil ( $ICC = .12$ ) als mangelhaft. Somit war die Expertenevaluation ungeeignet, um die Güte der Assistenzsystemkonzepte zu bestimmen. Aus diesem Grund entwickelte Krinner (2009), auf Basis der von den Experten genannten 119 Kriterien zur Konzeptgesamtbewertung, ein „Prozessmodell zur Evaluation von Assistenz“. Dieses, aus 15 Kategorien bestehende, Modell

umfasst die Konzeption, die Eigenschaften, die Implementation bis hin zum Einsatz eines Assistenzsystems.

#### **5.1.4 Mängel bei der Identifizierung der Leistungsdeterminanten**

In den Untersuchungen von Krinner (2009) konnten, trotz der augenscheinlichen Auslegungsvielfalt der konzipierten Assistenzsysteme, keine Faktoren gefunden werden, welche sich eindeutig auf die Konzeption und das Ergebnis von Assistenzsystemen auswirken.

Neben den Ergebnisbegründungen von Krinner (2009) sind weitere Ursachen denkbar. So könnte das Analyseraster der Taxonomie für Assistenzsysteme von Wandke (2005) und Wetzenstein (2004, s. Abschnitt 4.1.5) zu grob gestaltet sein, um alle relevanten Unterschiede zwischen den Konzepten zu identifizieren. Eine unangemessene Operationalisierung einer aV führt zu einer verminderten Konstruktvalidität (Sarris, 1990), wodurch wiederum keine sicheren Aussagen aus den Ergebnissen hinsichtlich der Forschungshypothesen abgeleitet werden können (Bainbridge, 1987; Sarodnick & Brau, 2006; Wandke, 2005). Unabhängig davon ermöglicht diese Taxonomie lediglich eine Unterscheidung hinsichtlich der Anzahl der berücksichtigten Dimensionsstufen. Es handelt sich um ein Ordnungssystem ohne Wertung der Art dieser Berücksichtigung. Somit kann eine qualitative Unterscheidung der Konzepte nicht vorgenommen werden.

Zudem könnte SAM 1.0 noch zu deterministisch sowie zu wenig komplex wirkend gewesen sein, um eine herausfordernde Konzeptionsaufgabe für die Entwickler anzubieten. Durch eine zu leichte Aufgabe können sämtliche Entwickler, unabhängig von den Stufen der uVn, alle zu lösenden Probleme identifizieren und alle notwendigen Assistenzsysteme konzipieren. (s. Nachtwei & Kain, 2010)

Ebenso erstaunt die mangelnde Beurteilerübereinstimmung zwischen den Evaluatoren nicht. Bei dem Vorgehen sind einige Aspekte kritisch anzumerken. Zum einen erhielten alle Experten nur eine kurze schriftliche Beschreibung von SAM, den MWBn und deren Aufgabe. Daher ist es fraglich, ob sich alle Experten dasselbe Bild von dem System und den damit verbundenen Anforderungen an die Assistenzsysteme aufbauen konnten. Zum anderen stammen die Experten aus verschiedenen Bereichen, d. h. deren Alltagsgeschäft umfasst die unterschiedlichsten Mensch-Technik-Interaktionssysteme zur Entwicklung und/oder Bewertung. Dadurch wurden von den Experten vermutlich verschiedene Bewertungsanker gesetzt respektive andere Referenzen verwendet (s. Tversky & Kahneman, 1974, zitiert nach Krinner, 2009). Des Weiteren ist anzunehmen, dass die Experten ein unterschiedliches Expertise-Niveau besaßen. Einige der Experten arbeiteten bereits seit vielen Jahren theoretisch und praktisch in Diesem Arbeitsfeld. Während andere hingegen wenig praktische Erfahrungen sammeln

konnten. Bisweilen existierte sogar nur eine rein theoretische Auseinandersetzung mit diesem Themengebiet im Rahmen einer Promotion.

## **5.2 ATEO Projektphase 2: Identifizierung der Leistungsdeterminanten**

Die bisherigen Untersuchungen im Projekt ATEO konnten kaum zur Identifizierung von Leistungsdeterminanten der Entwickler von Assistenzsystemen und Automaten sowie deren optimalen Ausprägungen beitragen. Daher beschloss die Autorin einen Schritt zurückzugehen, um eine systematische Literaturrecherche durchzuführen. Ziel war die Gewinnung eines Überblicks über die bisherigen Erkenntnisse aus der Forschung in Form eines deskriptiven Modells der Zusammenhänge zwischen Faktoren und der Leistung von Entwicklern von Automationen (s. Abschnitt 5.2.1). Aus diesem sollten relevante Variablen als uVn, aber auch als Kontrollvariablen für die experimentellen Untersuchungen im Projekt ATEO abgeleitet werden. Gleichzeitig wurde an den bisherigen Forschungsbemühungen aus der ersten ATEO-Projektphase angeknüpft. Dafür wurde die potentielle Leistungsdeterminante „Art der Information über das System“ einer systematischeren experimentellen Überprüfung unterzogen (s. Abschnitt 5.2.3). Nachfolgend wird diese als „Informationsverfügbarkeit“ bezeichnet. Zudem wurde eigens ein Instrument für die Bewertung der Güte der im Verlauf dieser Untersuchung gewonnenen Automatenkonzepte entwickelt (s. Abschnitt 5.2.2). Das Vorgehen und die Ergebnisse dieser drei Vorhaben werden nachfolgend dargestellt.

### **5.2.1 Modell der Leistungsdeterminanten von Entwicklern**

Im Zusammenhang mit der Leistung von Entwicklern von komplexen, dynamischen, technischen Systemen inklusive Automationen steht eine Vielzahl von Faktoren. Diese können u. a. in der Person des Entwicklers selbst, seinen Kollegen, seinen Vorgesetzten, der Aufgabe oder der Unternehmensstruktur begründet sein. Um diese Faktoren strukturiert, überblicksartig darzustellen, wurde von der Autorin gemeinsam mit der Diplomandin Rocholl (2010) das „deskriptive Modell der Leistungsdeterminanten von Entwicklern von komplexen, dynamischen, technischen Systemen“ entwickelt. Ausgangslage stellte die Durchführung einer Recherche psychologischer und ingenieurwissenschaftlicher Literatur dar, um ein bereits existierendes geeignetes Modell zu eruieren, welches sich mit den Leistungsdeterminanten von Entwicklern von Assistenzsystemen und Automaten zur Überwachung und Führung komplexer, dynamischer, technischer Systeme beschäftigt. Es konnten sieben Modelle identifiziert werden, welche die komplexen Zusammenhänge zwischen verschiedenen Faktoren und der Leistung von Entwicklern beschreiben. Bei diesen handelt es sich um die Modelle von Brown und Eisenhardt (1995), Ehrlenspiel (2003), Emmanuelides (1993), Frankenberger (1997), Gebert

(2004), Rosen (2005) und Verworn (2005). Jedoch bezieht sich keines der Modelle speziell nur auf Entwickler von Assistenzsystemen und Automaten. Einige beschäftigen sich zumindest mit Entwicklern von dynamischen, technischen Systemen (Rosen, 2005; Verworn, 2005). Zudem wurden nur die Modelle von Frankenberger (1997) und Verworn (2005) empirisch validiert. Es konnte somit kein adäquates Modell der Leistungsdeterminanten von Entwicklern von komplexen, dynamischen, technischen Systemen mit Automationen ausfindig gemacht werden.

Aus diesem Grund beschloss die Autorin ein eigenes zielgruppenbezogenes Modell zu entwickeln. Dafür wurde der Forschungsansatz der sogenannten „Performance Shaping Factors“ (PCF, s. Swain & Guttman, 1983) auf Entwickler übertragen. Hierbei handelt es sich um eine umfassende Sammlung von Faktoren, welche die Leistung von Operateuren bei der Prozessüberwachung und -führung beeinflussen.<sup>82</sup> Für die Ermittlung der Leistungsdeterminanten wurde die Literaturrecherche zu einer systematischen, intensiven qualitativen Metaanalyse<sup>83</sup> erweitert. Es sollten Variablen ermittelt werden, bei denen Zusammenhänge zu der Leistung von Entwicklern von Automationen empirisch nachgewiesen werden konnten. Diese Suche erwies sich allerdings als wenig ergiebig. Daher wurde sie auf empirische Arbeiten mit Entwicklern erweitert, deren Tätigkeit vergleichbar mit der von Entwicklern von Automationen ist. Hierbei handelt es sich bspw. um Tätigkeiten im Bereich der Software-entwicklung oder Produktentwicklung von komplexen, technischen Systemen. So konnten 453 Einzelergebnisse<sup>84</sup> gefunden werden. Diese stammen fast ausschließlich aus Untersuchungen der frühen Phasen der Entwicklung. Deshalb konnte das deskriptive Modell der Zusammenhänge zwischen Faktoren und der Leistung von Entwicklern nur für diesen Entwicklungszeitraum erstellt werden.<sup>85</sup> Auf Basis der zuvor erwähnten sieben Modelle und der Einzelergebnisse wurde ein Rahmenmodell erstellt. Wobei die *Hauptstruktur des Modells* an das *Input-Prozess-Output-Modell* (I-P-O-Modell) von Hackman (1987) angelehnt ist. Die *Input-Variablen* wurden den drei Hauptkomponenten „*Individuum*“, „*Team*“ und „*Rahmenbedingungen*“

---

<sup>82</sup> In den Modellen der 1980er Jahre thematisierten diese Faktoren nur das technische System und seine Umwelt. Ein neueres Modell, das Warnemünder Modell (Gérard, Huber, Nachtwei, Satriadarma, & Schubert, 2011), nimmt dagegen eine ganzheitliche Sicht auf diese Leistungsdeterminanten vor, indem es die Persönlichkeitsvariablen von Operateuren mit betrachtet.

<sup>83</sup> Die durchgeführte qualitative Metaanalyse stellt eine Kombination aus narrativen Review und quantitativer Metaanalyse dar (Rocholl, 2010). Es erfolgte zwar eine systematische und umfassende Literaturrecherche, doch wäre bei einer quantitativen Metaanalyse der forschungsrelevante Bereich enger eingegrenzt worden. Ebenso wurden die gefundenen Zusammenhänge in ein vergleichbares Datenniveau in Form des Effektgrößenmaßes  $r$  überführt. Allerdings wurden diese Effektgrößen nicht zu einem Wert verrechnet. Wie bei einem narrativen Review wurden die Untersuchungen und Modelle hinsichtlich deren wahrgenommener Stärken und Schwächen analysiert und beschrieben. Trotz konkreter Auswahlkriterien der im Modell einbeziehenden Zusammenhänge, liegen Unterschiede bei den Operationalisierungen der Variablen vor.

<sup>84</sup> Details zu den Variablen sind der Arbeit von Rocholl (2010) zu entnehmen.

<sup>85</sup> Ursprünglich war es das Ziel, für jede Phase des Entwicklungsprozesses das komplexe Wirkungsgefüge getrennt darzustellen.

zugeordnet. In die Hauptkomponente „Individuum“ wurden Variablen eingeordnet, die nur am Individuum erfasst werden können. Diese Hauptkomponente gliedert sich noch einmal in „Persönlichkeitseigenschaften“, „Fähigkeiten und Fertigkeiten“ sowie „Einstellungen und Werte“. Dagegen gingen in die Hauptkomponente „Team“ nur Variablen ein, deren Ausprägungen sich aus dem Team zusammensetzen. Zur Gruppierung dieser Variablen wurden die Kategorien „Teamzusammensetzung“, „Teamorganisation“ und „Teamklima“ gebildet. Variablen, welche das Unternehmen betreffen, bilden die Hauptkomponente „Rahmenbedingungen“. Diese beinhaltet die Kategorien „Führung“, „Organisation“ und „Aufgabe“. Die Zusammensetzung einer Organisation aus Teams, welche wiederum aus Individuen bestehen, wird durch die Verschachtelung dieser drei Hauptkomponenten im Modell veranschaulicht. Die **„Prozessvariablen“** sind alle Variablen, welche ein *Verhalten*, ein *Ereignis* oder einen sich *verändernden Zustand* beschreiben. Die **„Output-Variablen“** stellen die Leistung der Entwickler dar. Wie sich in der Literatur zeigt, existieren unterschiedlichste Operationalisierungen dieser Leistung.<sup>86</sup> Daher wird in dem Modell der Output durch die Leistungskriterien **„Teamleistung“**, **„Zeit“**, **„Kosten“** und **„Qualität des Produkts“** definiert. Zusätzlich wird im Modell die Hauptkategorie **„Umwelt“** abgebildet. In diese gehören Variablen, welche sich potentiell auf alle Unternehmen beziehen. Sie lassen sich einordnen in die Kategorien **„Markt“**, **„Kunde“**, **„Politik“** und **„Recht“**. Es konnten jedoch keine Zusammenhänge zu den anderen Hauptkategorien des Modells identifiziert werden. Vermutlich spielten Umweltvariablen, aufgrund ihrer schweren Kontrollierbarkeit (s. Emmanuelides, 1993), in der bisherigen Forschung nur eine untergeordnete Rolle.

Die gefundenen ***Einzelergebnisse*** stammen aus unterschiedlichen Untersuchungen mit verschiedenen Untersuchungsdesigns, Stichproben, Operationalisierungen und Erhebungsmethoden der einzelnen Variablen, statistischen Analysen, Teststärken usw. Deshalb wurden diese in das ***Effektgrößenmaß r*** umgewandelt. Dadurch können neben den Richtungen (positiv oder negativ) auch die Größen (kleine, mittlere oder große Effektgröße)<sup>87</sup> der Zusammenhänge auf vergleichbare und verständliche Weise in dem Modell dargestellt werden. Aufgrund von teilweise unzureichenden Ergebnispräsentationen in der Literatur ließ sich nur für 167 Einzelergebnisse die Effektgröße berechnen. Diese konnten wiederum in 58 Faktoren zusammengefasst werden, welche in einem Rahmenmodell abgebildet wurden (s. Anhang E).<sup>88</sup>

---

<sup>86</sup> Eine gelungene Übersicht über verschiedenste Operationalisierungen der Leistung von Entwicklern liefert Rocholl (2010).

<sup>87</sup> Konventionen für die Effektgröße  $r$  nach Cohen (1988, zitiert nach Sedlmeier & Renkewitz, 2008):  $r = 0.1$  ist ein kleiner,  $r = 0.3$  ist ein mittlerer und  $r = 0.5$  ist ein großer Effekt.

<sup>88</sup> Um die restlichen 286 interessanten Forschungsbefunde nicht unbehandelt zu lassen, wurde eine Erweiterung des Modells vorgenommen (s. Rocholl, 2010).

Aufgrund des zeitlichen Vorkommens der einzelnen Variablen im Entwicklungsprozess liegt dem Modell die Annahme zu Grunde, dass die Input- und Prozessvariablen die Leistung von Entwicklern beeinflussen. Anzunehmen ist zudem, dass die Input-Variablen auf die Prozessvariablen einwirken. Daher werden Kausalzusammenhänge vermutet, welche jedoch nicht durch einen Korrelationskoeffizienten  $r^{89}$  dargestellt werden können (Bortz, 1999). Zusätzlich sind Feedbackschleifen im Modell berücksichtigt worden. Die Autorin hat die erste veröffentlichte Modellversion (Rocholl, 2010), welche nur eine Feedbackschleife zwischen Output und Input umfasste, um Feedbackschleifen zwischen Output und Prozess, Prozess und Input und innerhalb des Prozesses selbst erweitert. Allerdings ist dabei zu beachten, dass bei den Input-Variablen einige der individuellen und somit auch Gruppen-Variablen teilweise nicht (bspw. Persönlichkeitseigenschaften) bzw. nur über einen längeren Zeitraum und schwer (Einstellungen und Werte) veränderbar sind.

Die Gültigkeit des Modells wurde gemeinsam mit zehn Experten aus dem Bereich Software- oder technischer Produktentwicklung in einem Workshop validiert. Insgesamt schrieben die Experten dem Modell eine hohe Passung zu deren persönlichen Erfahrungen aus. Einige Experten wünschten sich die Verwendung eines einheitlichen Abstraktionsniveaus für die Faktoren. Dies würde jedoch eine Zusammenlegung von Faktoren und damit Informationsverlust mit sich bringen. Das Modell soll im Projekt ATEO und darüber hinaus als Informationsbasis dienen, welche Variablen als potentielle Leistungsdeterminanten in Untersuchungen mit Entwicklern von Automationen systematisch überprüft, mit erhoben oder ausgeschaltet werden sollten. Dadurch soll nach und nach ein experimentell validiertes Modell der Leistungsdeterminanten speziell nur für diese Entwickler entstehen. Deshalb wurde der Informationsgehalt des Modells beibehalten.

#### **5.2.1.1 Modell Output-Operationalisierung „Teamleistung“**

Für einige Variablen konnten Zusammenhänge nur mit einzelnen Output-Variablen aufgefunden werden. Daher wurde zusätzlich zum Gesamtmodell speziell für die Operationalisierungen „Teamleistung“, „Produktqualität“, „Zeit“, „Effizienz“ und „Veränderbarkeit“ jeweils ein eigenes Modell erstellt (s. Rocholl, 2010). Nachfolgend wird auf das Modell mit der Output-Variable „Teamleistung“ etwas näher eingegangen (s. Abbildung 22), da es sich bei der Teamleistung um die in der Literatur am häufigsten wiedergegebene Operationalisierung der

---

<sup>89</sup> Ein Korrelationskoeffizient lässt immer vier Interpretationsmöglichkeiten zu. Es können beide Variablen jeweils die andere Variable oder sich gegenseitig kausal beeinflussen haben. Ebenso kann eine weitere oder mehrere andere Variablen einen kausalen Einfluss auf eine oder beide Variablen haben. (Bortz, 1999)

Leistung von Entwicklern handelt (Rocholl, 2010). Wie dem Modell zu entnehmen ist, stehen u. a. die Expertise und das Leistungsmotiv der einzelnen Entwickler in einem positiven Zusammenhang mit der Teamleistung. Negative Zusammenhänge finden sich bspw. bei heterogen zusammengesetzten Teams hinsichtlich der drei Big Five Dimensionen „Extraversion“, „Gewissenhaftigkeit“ sowie „Verträglichkeit“ (s. Goldberg, 1981) und ihrer Teamleistung. Bei den Prozessvariablen weist u. a. die „Anstrengung“ und die „interne Kommunikation“ einen positiven Zusammenhang zur Teamleistung auf. Die Bedeutsamkeit einiger Faktoren dieses Modells für die Leistung von Entwicklern von Automationen wurde durch die Autorin im Rahmen der Untersuchung „Informationsverfügbarkeit“ validiert (s. Abschnitt 5.2.3).

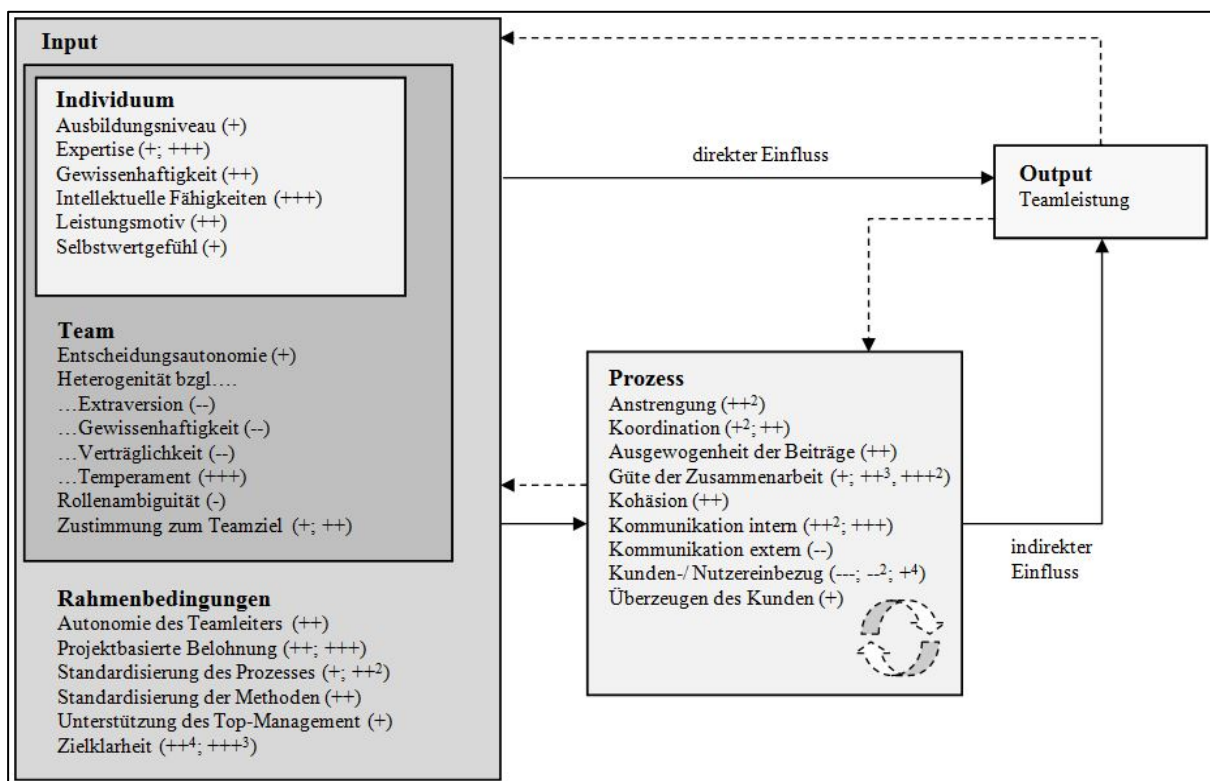


Abbildung 22: „Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme“ mit der Output-Variable „Teamleistung“ erweitert um Feedbackschleifen (Original s. Rocholl, 2010)

*Anmerkungen.* In dem Modell werden die Richtung (+ = positiv und - = negativ) und die Stärke (+/- = kleine, ++/- = mittlere und +++/- = große Effektgröße) eines Zusammenhangs einzelner Variablen dargestellt. Dabei ist zu beachten, dass die über die durchgehenden Linien dargestellten Kausalzusammenhänge nur aufgrund des zeitlichen Vorkommens der Variablen angenommen sind. Gestrichelte Linien kennzeichnen, dass Prozess- oder Output-Variablen zurückwirken können.

### 5.2.2 Bewertungsinstrument der Güte von Automatikkonzepten

In der Untersuchung der Autorin zur systematischen Überprüfung der Leistungsdeterminante „Informationsverfügbarkeit“ mussten Entwickler Automaten für die Überwachung und Führung von SAM konzipieren (s. Abschnitt 5.2.3). Als aV wurde u. a. die „Güte der Automatikkonzepte“ verwendet. Um Güteunterschiede zwischen diesen Konzepten ermitteln zu können, mussten diese einer umfassenden Analyse unterzogen werden. Wie sich bereits in der ersten ATEO-Projektphase zeigte, genügte der, von Krinner (2009) selbst entwickelte, „Fragebogen zur Bewertung von Assistenzkonzepten“ diesem hoch auflösenden Anspruch nicht. Daher leitete Krinner (2009) aus diesem Fragebogen sowie den zusätzlich in der Expertenevaluation gewonnenen Beurteilungskriterien ein „Prozessmodell zur Evaluation von Assistenz“ ab. Es listet 15 relevante Kategorien für die Bewertung von Automationskonzepten auf. Allerdings sind diese Kategorien noch zu grob und ungenau für eine Gütebewertung der Automatikkonzepte formuliert.

In der Literatur, in Normen, Richtlinien und Heuristiken zur Entwicklung, Gestaltung und Bewertung technischer Systeme konnte die Autorin ebenfalls kein geeignetes Instrument zur Gütebewertung der Automatikkonzepte auffinden. Die identifizierten Bewertungskriterien waren entweder zu allgemein oder zu spezifisch.

Aus diesem Grund wurde durch die Autorin eine theoriegeleitete, umfassende, systematische Bewertung der Automatikkonzepte in Anlehnung an eine qualitative Inhaltsanalyse mit skalierender Strukturierung nach Mayring (2000, 2007) durchgeführt. Für die Entwicklung des Bewertungsinstruments wurden von den Instrumenten und den 119 Beurteilungskriterien der Experten von Krinner (2009) sowie aus der Literatur, den Normen, den Richtlinien und den Heuristiken die Kriterien verwendet, welche relevant für die Bewertung von Automaten sind. Aus der Außensicht auf SAM überwachen und führen Automaten im Sinne eines vollautomatischen Systems anstelle von Operateuren den Prozess von SAM.<sup>90</sup> Zusätzlich wurden Kriterien berücksichtigt, welche bedeutsam für das kooperative Tracking sowie für die Bewertung von Mensch-Maschine-Schnittstellen und Assistenzsystemen sind. Diese Kriterien konnten herangezogen werden, da die MWB reale Personen sind. Aus ihrer Sicht handelt es sich bei den Automaten um Assistenzsysteme, durch welche sie genauso wie durch einen Operateur bei der Ausführung der Tracking-Aufgabe unterstützt werden.<sup>91</sup> Somit

---

<sup>90</sup> Eine kooperative Automation zur Überwachung und Führung von SAM stand erst in der dritten ATEO-Projektphase im Forschungsfokus (s. Niestroj, 2014; von Bernstorff, 2014). Daher wurden Kriterien zur Bewertung dieser Automationsform nicht mit einbezogen.

<sup>91</sup> Dabei können die MWB nicht unterscheiden, ob eine Unterstützung durch einen Operateur oder eine Automatik erfolgt.



ist es möglich, Automatikkonzepte für die Überwachung und Führung von SAM aus zwei Perspektiven heraus zu bewerten. Zusätzlich flossen die Ergebnisse der HAA für die OA-S (s. Anhang C.1.2) und für die AA-S (s. Anhang C.1.1) in das Bewertungsinstrument mit ein. Außerdem wurden die vorgegebenen Einschränkungen für die Automatikkonzeption im Bewertungsinstrument berücksichtigt. Beispielsweise sollten die MWB durch die Automaten keinerlei Kenntnis über ihre gegensätzlichen Instruktionen erhalten (s. Anhang G.1.1).

Durch dieses Vorgehen konnte deduktiv ein fundiertes Bewertungsinstrument entwickelt werden, welches einer objektivierten Gütebewertung von Automatikkonzepten für SAM dienen soll (s. Sarodnick & Brau, 2006). Um eine hohe Differenzierung zwischen den Automatikkonzepten erreichen zu können, wurde der Verlust einer praktischen Handhabbarkeit des Bewertungsinstruments in Kauf genommen. So entstand ein „**Bewertungsinstrument der Güte von Automatikkonzepten**“, welches aus zwei Komponenten besteht.<sup>92</sup>

Die eine Komponente ist eine **Bewertungsfunktionsmatrix**, die sämtliche relevanten Funktionen von Automaten zur Überwachung und Führung von SAM auflistet (s. Abschnitt 5.2.2.1). Zudem werden für jede Funktion die verschiedenen Auslegungsmöglichkeiten aufgeführt und durch Skalenwerte gewichtet. Der Begriff „Auslegung“ ist in der Entwicklungspraxis ein verbreiteter Terminus. In Bezug auf technische Systeme ist neben der Materialfestlegung die konkrete funktionale, ergonomische und gebrauchstaugliche Gestaltung des Systems zur Erfüllung seines Ziels gemeint. So finden sich z. B. im Bereich der Fahrzeugführung für bremsende Fahrerassistenzsysteme unterschiedliche Auslegungen, die von ihrer jeweiligen Aufgabe abhängen. Eine Bremsung zur Anpassung des Abstands, aufgrund eines langsamer werdenden Vorderfahrzeugs, erfolgt beispielsweise für die Fahrzeuginsassen als komfortable wahrnehmbar, im Rahmen einer automatischen Distanzregelung. Dagegen führt ein Notbremsystem zur Vermeidung einer Kollision mit einem Vorderfahrzeug in einer zeitkritischen Situation eine für die Insassen unkomfortable Vollverzögerung aus. In SAM existieren für die Reduzierung der vertikalen Objektgeschwindigkeit ebenfalls verschiedene Auslegungsmöglichkeiten. So kann die vertikale Objektgeschwindigkeit nur für einen kurzen Moment bis hin zum Stillstand – leicht oder stark, graduell oder abrupt – reduziert werden.

Die zweite Komponente des Bewertungsinstruments ist ein **Bewertungsfragebogen**. In diesem sind weitere relevante Kriterien in Form von Items mit Antwortskala aufgelistet, welche aus Gründen der Übersicht nicht in die Funktionsmatrix integriert wurden (s. Abschnitt 5.2.2.2).

---

<sup>92</sup> Das vollständige Instrument kann bei der Autorin angefragt werden.

Beide Instrumentkomponenten wurden in mehreren iterativen Schritten an den Automatikkonzepten selbst hinsichtlich der Angemessenheit für eine umfassende sowie differenzierende Gütebewertung überprüft und ggf. angepasst. Für die Funktionsmatrix konnten auf diese Art und Weise zudem sinnrekonstruktiv weitere relevante Funktionen identifiziert werden. Bevor diese in die Funktionsmatrix integriert wurden, wurde deren Relevanz für die Überwachung und Führung von SAM von der Autorin im ATEO-Team diskutiert. Deren Auslegungsmöglichkeiten samt Gewichtung wurden wiederum theoriegeleitet bestimmt. Beispielsweise erhalten durch die Funktionsmatrix nur Funktionen zur permanenten Vermittlung einer Information die höchste Punktzahl, für welche die visuelle Modalität vorgesehen ist. Eine permanente auditive Informationsvermittlung wäre zu ablenkend und störend für die MWB (s. Campbell, Richard, Brown, & McCallum, 2007). Daher erhalten Funktionen mit dieser Art von Informationsvermittlung die geringste Punktzahl. Ebenso wird an Funktionen mit visuellen Informationsvermittlungsarten mit unspezifischen Charakter, wie bspw. Licht (Herczeg, 1994, 2009), die geringste Punktzahl vergeben, wenn eine explizite Information – etwa die Güte der Objektregelung durch die MWB – permanent vermittelt werden soll.

Niestroj (2014) ermittelte für die Funktionsmatrix eine Inter-Rater-Reliabilität mit einem durchschnittlichen Cohens Kappa-Koeffizienten von  $.40^{93, 94}$  ( $SD = .06$ ). Grundlage hierfür bildeten die Bewertungen von zehn Automatikkonzepten<sup>95</sup>, die zum einen durch Niestroj (2014) selbst gemeinsam mit einer studentischen Hilfskraft und zum anderen durch die Autorin vorgenommen wurden. Die Autorin versuchte, bei der Bewertung der Konzepte Interpretationen zu vermeiden. Niestroj (2014) verfolgte dahingegen den Ansatz, Lücken durch Interpretationen zu schließen und somit nicht allein die Leistung der Konzeptentwickler zu bewerten. Unter anderem deshalb lässt die ermittelte Beurteilerkonkordanz diese Instrumentkomponente ausreichend objektiv zur Bewertung der Güte von Automatikkonzepten erscheinen. Eine ausführlichere Begründung hierfür findet sich im Anhang F.2.1.

Beide Instrumentkomponenten werden nachfolgend vorgestellt.

---

<sup>93</sup> Der Wertebereich des Kappa-Koeffizienten von Cohen liegt zwischen  $-1 \leq \kappa \leq 1$ , wobei ein Wert von -1 keine, von 1 eine völlige und von 0 eine zufällige Übereinstimmung bedeuten (Bortz & Lienert, 1998). Der Wert .40 markiert die obere Grenze für die Kategorie „ausreichende Übereinstimmung“ (Landis & Koch, 1977).

<sup>94</sup> Bei der Kodierung der konkreten Gewichtungspunktwerte mit 1 ergab sich ein mittlerer Cohens Kappa-Koeffizient von .41 ( $SD = .05$ ) (Niestroj, 2014). Dieser Wert entspricht der unteren Grenze der Kategorie „mittelmäßige Übereinstimmung“ (Landis & Koch, 1977).

<sup>95</sup> Diese Konzepte entstanden im Rahmen der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“ in der dritten Bedingung der uV (s. Abschnitt 5.2.3).

### 5.2.2.1 Bewertungsfunktionsmatrix

Die Funktionsmatrix besteht aus 95 Funktionen mit jeweils unterschiedlichen Funktionsaspekten. Aus diesem Grund wird hier nachstehend lediglich die Struktur mit einigen Beispielen und den ihr zugrundeliegenden Erkenntnissen aus Forschung und Praxis vorgestellt.

Ein technisches System muss alle notwendigen Funktionen für die Ausführung der zu bewältigenden Aufgaben aufweisen (Sarodnick & Brau, 2006). Bezogen auf die kooperative Tracking-Aufgabe erweisen sich jene Automatikfunktionen als relevant, welche die MWB kontinuierlich bei der genauen und schnellen Objektregelung unterstützen. Exemplarisch hierfür ist eine permanente Vorausschau des Pfadverlaufs, die mit einer Streckenvorausschau einer Rennspielsimulation vergleichbar ist. Da die einzelnen Pfadelemente spezielle Anforderungen an die MWB stellen, sind zusätzlich Funktionen notwendig, welche die MWB bei der Objektregelung nur an bestimmten Pfadelementen unterstützen. Beispielhaft hierfür ist eine Funktion zur Richtungsempfehlung an einer Gabelung. Solch eine Funktion ist vergleichbar mit der Richtungsempfehlung an einer Kreuzung durch ein Navigationssystem im Fahrzeug. Um die Auflistung der einzelnen, teilweise pfadspezifischen Funktionen in der Matrix übersichtlich zu gestalten, wurde die Hauptkategorie „Pfadelement“ verwendet (s. Tabelle 21). In ihr wird neben dem gesamten Pfad zusätzlich zwischen den sechs Pfadelementen „Gerade“, „starke Kurve“, „Gabelung“, „statisches Hindernis 25 %“, „statisches Hindernis 50 %“ und „dynamisches Hindernis“ unterschieden. Zudem sind in dieser Hauptkategorie die Pfadelemente „statisches Hindernis allgemein“ und „Hindernis allgemein“ aufgelistet. Dadurch kann bewertet werden, ob in einem Konzept die spezifischen Anforderungen der einzelnen Hindernisarten, nur der statischen Hindernisse allgemein und des dynamischen Hindernisses oder lediglich für Hindernisse allgemein berücksichtigt sind.

Einige der pfadelementspezifischen Funktionen müssen aktiv werden, bevor das entsprechende Pfadelement im Sichtbereich der MWB erscheint, damit ihnen genügend Zeit für eine adäquate Objektregelung verbleibt. Ein Beispiel hierfür ist eine Funktion, welche die MWB auf ein vorausliegendes Hindernis hinweist. Eine vergleichbare Funktionalität aus dem Bereich der Fahrzeugführung stellt die Anzeige von vorausliegenden Baustellen durch ein Navigationssystem dar. Aufgrund der unterschiedlichen Aktivitätsbereiche der pfadelementspezifischen Funktionen wurde in der Hauptkategorie „Pfadelement“ zusätzlich für die Pfadelemente „scharfe Kurve“, „Hindernisarten“ und „Gabelung“ zwischen „kurz vor“ und „im Sichtbereich“ der MWB differenziert (s. Tabelle 21 und detailliert im Anhang F.2.2).

Auf einem Pfadelement können unterschiedliche Situationen eintreten, in denen die MWB bei ihrer Tracking-Aufgabe pfadelements- und situationsspezifisch durch verschiedene

Tabelle 21: Struktur der Bewertungsfunktionsmatrix mit Darstellung der Hauptkategorie „Funktionsklasse“ sowie einzelner Kategorien von deren Unterkategorie, der Hauptkategorie „Pfadelement“ und einzelner Kategorien von deren Unterkategorie „Situation“ sowie exemplarischer Funktionen. Ein X kennzeichnet, für welche Situation eine Funktion geplant ist.

		Hauptkategorie „Pfadelement“	gesamter Pfad			...	sobald dynamisches Hindernis im Sichtbereich bis Objekt Hindernis passiert hat		
		Unterkategorie „Situation“	allg.	...	MWB können Pfad-verlassen nicht mehr abwenden	...	allg.	...	Gefahr Hindernis-kollision, MWB können dies noch abwenden
Hauptkategorie „Funktions-klasse“	Unter-kategorie der Funktions-klasse	Funktion							
Informations-akquise	Daten-erfassung	Pfadmittel-linienverlauf	X						
		...							
	Informations-vermittlung	Pfadvoraus-schau	X						
		...							
Informations-analyse	Daten-berechnung	Kollisions-gefahr mit Hindernis					X		
		...							
	Informations-vermittlung	Warnung Gefahr der Kollision							X
		...							
Entscheidungs-findung	Entscheidung	Bewertung und Selektion der Geschwindigkeit des Objekts bei dynamischen Hindernis					X		
		...							
	Informations-vermittlung	Anzeige horizontale Ideal-Joystick-Eingabe MWB 1 & 2	X						
		...							
Handlungs-ausführung	Veränderung des Anteils an der Objekt-regelung zwischen den MWBn	Veränderung des vertikalen Anteils an der Objektregelung zwischen den MWBn	X						
		...							
	...	...							
	Eingriff in die Objekt-regelung	Übernahme der Regelung der Geschwindig-keit des Objekts allg.			X				
		...							

Funktionen unterstützt werden müssen. Beispielsweise kann bei einem Hindernis die Gefahr bestehen, dass das Tracking-Objekt mit ihm kollidiert. In diesem Fall kann für die MWB eine Funktion hilfreich sein, welche sie vor einer möglichen Kollision warnt. Ebenso kann es sein, dass die MWB das Tracking-Objekt zu großzügig um das Hindernis regeln und dabei das Verlassen des Pfads selbst nicht mehr abwenden können. Dann wäre eine Funktion sinnvoll, die direkt in den MWB-Objekt-Tracking-Umgebung-Regelkreis eingreift, um die entstandene kritische Situation schnell zu deeskalieren. Die im jeweiligen Pfadelement identifizierten relevanten Situationen bilden deshalb die Unterkategorie „Situation“ der Hauptkategorie „Pfadelement“ (s. Tabelle 21 sowie Details im Anhang F.2.2). Dadurch kann eine übersichtliche Gestaltung der Funktionsmatrix erreicht werden.

Zusätzlich wurde zum Erhalt einer übersichtlichen Struktur der Funktionsmatrix die Hauptkategorie „Funktionsklasse“ mit den Kategorien „Informationsakquise“, „Informationsanalyse“, „Entscheidungsfindung“ sowie „Handlungsausführung“ (s. Tabelle 21), angelehnt an das Modell der Automatisierungsstufen von Parasuraman et al. (2000), verwendet (s. Abschnitt 4.1.5).

Bei der Identifizierung und Gewichtung der relevanten Automatikfunktionen sowie deren Auslegungsmöglichkeiten wurden die mit den einzelnen Automationsausmaßen verbundenen Probleme berücksichtigt (s. Abschnitt 4.1.2). Hierbei wurde den Empfehlungen von Parasuraman (2000) und et al. (2000) gefolgt (s. Abschnitt 4.1.5 und 4.1.6). In der Funktionsmatrix werden deshalb für die Informationsakquise und -analyse vollautomatisierte Funktionen vorgesehen. Dagegen sind für die Funktionen zur Unterstützung der Entscheidungsfindung und Handlungsausführung der MWB vorrangig mittlere und nur in Notfallsituationen, wenn die MWB eine Situation nicht mehr von alleine deeskalieren können, hohe Stufen der Automation vorgesehen. Außerdem wurde sich bei der Identifizierung der relevanten Funktionen und deren Auslegungsmöglichkeiten an den weiteren Taxonomien aus dem Abschnitt 4.1.5 und aus dem Bereich der Fahrzeugführung orientiert. Letztere eignen sich besonders aufgrund der Charakteristik der Tracking-Aufgabe, welche als eine sehr rudimentäre Fahrzeugregelaufgabe angesehen werden kann (s. Burmester, Graf, Hellbrück, & Meroth, 2008). Die Automatisierungsstufen dieser bereichsspezifischen Klassifikationsansätze weisen hohe Ähnlichkeiten auf. So wird zumeist zwischen Information, Hinweis, Warnung, Handlungsempfehlung und Eingriff differenziert, wenngleich diese teilweise auch andere Bezeichnungen tragen (s. Anhang F.1 oder Baber & Wankling, 1992; Cotter et al., 2006; Frank & Reichelt, 2001; N. Fricke, De Filippis, & Thüning, 2008; Gründl, 2005; Lermer, 2010; Thoma, 2010). Außerdem

wird in der Funktionsmatrix der Forderung nachgekommen, die einzelnen Stufen der Automation in Abhängigkeit der Kritikalität der Situation respektive der Dringlichkeit einer menschlichen Handlung einzusetzen (Moray et al., 2000).

Ferner konnten durch die Bewertungskriterien der Experten von Krinner (2009) relevante Funktionen identifiziert werden (s. Anhang F.2.3).

#### 5.2.2.1.1 Hauptkategorie „Funktionsklasse“

Nachfolgend wird näher auf die Hauptkategorie „Funktionsklasse“ mit einigen beispielhaften Funktionen eingegangen. Um deren Verständlichkeit zu erhöhen, werden teilweise vergleichbare Automationen von realen MMS dargestellt.

##### 5.2.2.1.1.1 Funktionen der Kategorie „Informationsakquise“

Die Kategorie „Informationsakquise“ setzt sich aus den beiden Unterkategorien „Datenerfassung“ und „Informationsvermittlung“ zusammen. Wenn eine Automatik die MWB adäquat unterstützen soll, muss sie permanent den Zustand der Tracking-Umgebung, des Tracking-Objekts, der beiden MWB und anderer Automaten kennen. Die dafür notwendigen Informationen werden durch die Funktionen der Unterkategorie „**Datenerfassung**“ erhoben. Beispielhaft sind Funktionen zur Erfassung der Geschwindigkeit des dynamischen Hindernisses oder der Reaktionszeiten der MWB. In realen technischen Systemen werden Umwelt und Benutzerdaten bspw. durch Kameras, Radar-, Druck- und Vereisungssensoren erfasst.

Einige der erfassten Daten können den MWBn als reine *Information* dargeboten werden, um deren Situationsbewusstsein zu erhöhen. Funktionen, welche solch eine Aufgabe innehaben, sind der Unterkategorie „**Informationsvermittlung**“ zugeordnet. Diese Funktionen können immer aktiv sein. Sie können z. B. permanent die Ist-Joystick-Eingabe je MWB anzeigen, bspw. in einem Koordinatensystem positioniert neben dem Tracking-Pfad nahe<sup>96</sup> dem Tracking-Objekt. Ebenso können solche Funktionen ereignisabhängig aktiv sein, etwa zur Anzeige der Aktivität einer anderen Automatikfunktion. Diese andere Funktion kann bspw. im Kurvenbereich die vertikale Objektgeschwindigkeit reduzieren, um ein Verlassen des Pfads aufgrund zu hoher Objektgeschwindigkeit zu verhindern. Solch eine automatisierte Aktivität

---

<sup>96</sup> Relevante visuelle Informationen sollten nah beieinander positioniert werden (Vöhringer-Kuhnt, 2010). Auf den Ort, an dem viele relevante Informationen platziert sind, richtet sich die menschliche Aufmerksamkeit (Wickens, 2007). Ist der perzeptive Aufwand aufgrund zu großer Abstände zwischen relevanten Informationen zu hoch, findet keine Informationsaufnahme statt (Sarodnick & Brau, 2006; Wickens, 2007). Daher sollten foveal aufzunehmende Informationen nah am zu regelnden Objekt positioniert sein, um den Aufwand gering zu halten und die Aufmerksamkeit nicht von diesem abzuziehen.

muss den MWBn innerhalb eines adäquaten Zeitraums transparent gemacht werden (s. DIN ISO/IEC 12199, 1995; Billings, 1997; Nielsen, 1994), damit sie nicht vom Objektverhalten sowie dem Zustandswechsel der direkt in den MWB-Objekt-Tracking-Umgebung-Regelkreis eingreifenden Funktion überrascht werden (s. Preim, 1999). Durch eine Aktivitätsanzeige einer Funktion werden die MWB in die Lage versetzt, den Zustand des Tracking-Objekts in Bezug zu den eigenen Handlungen sowie anderen Ereignissen zu setzen und somit im Laufe der Zeit deren Auswirkungen zu antizipieren (s. Herczeg, 2009). Solch eine Informationsgabe führt in einem MMS, an dem ein oder mehrere Menschen beteiligt sind, zu einer Verbesserung der Gesamtleistung (Krahl, LoVerde, & Scerbo, 1999, zitiert nach Kaber et al., 2001). Dabei sollte ebenso vermittelt werden, welche konkrete Aktivität eine Funktion aktuell ausführt. Nur so kann eine Automation den Menschen entlasten (Maier, 2014). In realen MMS werden Aktivitätsanzeigen automatisierter Funktionen vielfach umgesetzt. So wird z. B. dem Fahrer im Cockpit von manchen Motorrädern die Reduktion des Motordrehmoments durch die Traktionskontrolle zur Vermeidung eines durchdrehenden oder wegrutschenden Hinterrades angezeigt.

In die Funktionsmatrix wurden nur Funktionen aufgenommen, welche die Aktivität von direkt in den MWB-Objekt-Tracking-Umgebung-Regelkreis eingreifenden Funktionen an die MWB vermitteln. Die Aktivität von anderen Funktionen an die MWB zu vermitteln, könnte zu einer zu hohen Anzahl an Aktivitätsanzeigen führen. Damit einher können eine Informationsüberflutung (s. Sarter et al., 1997; Sheridan, 2000) und zu hohe Beanspruchung der MWB gehen (s. Parasuraman, 2000).

In die Unterkategorie „Informationsvermittlung“ fallen zudem Funktionen, welche den MWBn *Hinweise* auf Pfadelemente geben, kurz bevor diese im Sichtbereich der MWB erscheinen. Hierbei handelt es sich um Hinweise, etwa um auf eine vorausliegende starke Kurve oder Gabelung aufmerksam zu machen, bspw. in Form eines Icons des jeweiligen Pfadelements. Hinweise sollten in der Nähe des Tracking-Objekts platziert sein, ohne jedoch den Pfad abzudecken. Vergleichbare Funktionalitäten bieten sogenannte Effizienzassistenten im Bereich der Fahrzeugführung an. Diese zeigen dem Fahrer z. B. vorausliegende Kurven oder Kreuzungen, teilweise grafisch veranschaulicht, im Bereich des Kombiinstrumentes oder dem Head-Up-Display an, damit dieser die Fahrzeuggeschwindigkeit entsprechend anpassen kann.

Im Allgemeinen liefert ein Hinweis frühzeitig wichtige zusätzliche Informationen über eine Situation, welche dem Menschen helfen, diese leichter zu interpretieren und vorausschauend sowie souveräner zu bewältigen (Maier, 2014; Thoma, 2010). Ein Hinweis soll auf eine potentielle Gefahrensituation aufmerksam machen (Wilson, Butler, McGehee, & Dingus, 1997). Dadurch kann der Mensch bereits unter nicht zeitkritischen Bedingungen eine Gefahr

vermeiden (Thoma, 2010), wobei ihm ein größerer Handlungsspielraum zur Verfügung steht (Lerner, Kotwal, Lyons, & Gardner-Bonneau, 1996). Außerdem fällt seine Reaktion aufgrund der geringeren Kritikalität weniger drastisch aus (Maier, 2014), wodurch wiederum die Gefahr von negativen Handlungskonsequenzen verringert ist. Zusätzlich führt ein Hinweis über eine zukünftige Situation zur kognitiven Entlastung beim Menschen (Wickens et al., 2004).

#### **5.2.2.1.1.2 Funktionen der Kategorie „Informationsanalyse“**

Die Kategorie „Informationsanalyse“ teilt sich auf in die Unterkategorien „Datenberechnung“ und wiederum „Informationsvermittlung“.

Die Unterkategorie **„Datenberechnung“** bildet sich aus Funktionen, welche aus den erfassten Daten der Funktionen der „Informationsakquise“ neue Variablen berechnen. Beispielfähig wäre die Bestimmung der Güte der Joystick-Eingaben der MWB durch einen Vergleich der Ist- mit den Ideal-Joystick-Eingaben. Ebenso gehören in diese Kategorie Funktionen, welche die erhobenen Tracking-Daten dahingehend analysieren, ob sich das Tracking-Objekt in einer aufgabenangemessenen oder in einer kritischen Situation befindet, z. B. kurz vor dem Verlassen des Tracking-Pfads.

Die Unterkategorie **„Informationsvermittlung“** beinhaltet Funktionen, welche einige der neu berechneten Variablen den MWBn zur Erweiterung ihres Situationsbewusstseins darbieten. Eine solche Informationsvermittlung stellt das Feedback zur Objektregelungsgüte an die MWB dar. Das Wissen über die eigene Objektregelungsgüte ist für die MWB auch in unkritischen Situationen nützlich. Daher sollte es permanent vermittelt werden. Der Mensch benötigt zum Erlernen und Beherrschen eines technischen Systems ein Feedback zur Angemessenheit der eigenen Handlungen (Bainbridge, 1983). Dadurch kann er den Systemzustand zu den eigenen Handlungen in Beziehung setzen (Kaber et al., 2001). Hierbei sollte der Zusammenhang zwischen Handlung und Fehler explizit dargestellt werden (DIN EN ISO 9241-110, 2006; Sarodnick & Brau, 2006). Durch eine solche Informationsvermittlung können die MWB die von ihnen selbst wahrgenommene Genauigkeit (Position Tracking-Objekt in Relation zum Tracking-Pfad) und Geschwindigkeit (Tachoanzeige im Tracking-Objekt) des Tracking-Objekts leichter in Bezug zur eigenen Objektregelung setzen. Dadurch können sie ungeeignetes Tracking-Verhalten identifizieren und zukünftig vermeiden. Sowohl ein individuelles als auch ein Team-Feedback zu den eigenen Handlungen verbessern das Situationsbewusstsein des Menschen (Kaber & Endsley, 1998) sowie die MMS-Leistung (Krahl et al., 1999, zitiert nach Kaber et al., 2001). Ein ebenfalls aus dem Fahrzeugbereich gewähltes Beispiel für ein Feedback zur Verhaltensgüte sind die Benzin- oder Stromverbrauchsanzeigen, welche dem Fahrer ein unmittelbares Feedback zum Verbrauch aufgrund seiner aktuellen Fahrweise geben. Diese



existieren in verschiedenen Formen, von einfachen Zahlenwerten bis hin zu grafischen Visualisierungen, wie z. B. sich farblich ausfüllenden Blättern.

Zudem sind Funktionen Bestandteil dieser Unterkategorie, welche die MWB vor Gefahren warnen. Die MWB befinden sich zum Warnzeitpunkt in einer Situation mit hoher Kritikalität, in welcher sie die Gefahr jedoch noch selbst abwenden können. Hierbei kann es sich bspw. um eine **Warnung** vor einer Kollision mit einem Hindernis handeln. Im Bereich der Bahnfahrt bietet das Kollisionswarnsystem eine vergleichbare Funktionalität. Dieses warnt den Triebfahrzeugführer visuell und akustisch vor herannahenden Zügen auf demselben Gleis.

Eine Warnung erfolgt i. d. R. in einer zeit- und sicherheitskritischen Situation zur unmittelbaren Lenkung der Aufmerksamkeit auf die akute Gefahr (Lerner, 2010; Selcon, 1999; Thoma, 2010). Dafür benötigt eine Warnung eine hohe Signalentdeckungsrate (N. Fricke, 2009; Maier, 2014). Zugleich muss diese auf den Menschen alarmierend wirken (Edworthy & Adams, 1996; Selcon, 1999). Außerdem soll eine Warnung dem Menschen eine nützliche Interpretation der Situation, deren Kritikalität, die Dringlichkeit einer Handlung, konkrete Handlungsempfehlungen oder die Folge einer falschen Reaktion liefern (Edworthy & Adams, 1996; Lerner et al., 1996), um so negative Auswirkungen der Situation zu vermeiden oder zumindest zu vermindern (Laughery & Hammond, 1999; Thoma, 2010). Für die Deeskalation der Situation muss dem Menschen ausreichend Zeit zur Informationsverarbeitung, Handlungsentscheidung und -ausführung zur Verfügung stehen (Lee, McGehee, Brown, & Reyes, 2002; Thoma, 2010). Des Weiteren muss eine Warnung hoch intuitiv verständlich (N. Fricke, 2009) und verwechslungsarm gestaltet sein (Maier, 2014). Nur so kann der Mensch schnell eine situationsangemessene Entscheidung treffen und eine entsprechende Handlung zur Abwendung der Gefahrensituation ausführen (N. Fricke, 2009).

#### 5.2.2.1.1.3 Funktionen der Kategorie „Entscheidungsfindung“

Die Kategorie „Entscheidungsfindung“ setzt sich aus den beiden Unterkategorien „Entscheidung“ und abermals „Informationsvermittlung“ zusammen.

Die Funktionen der Unterkategorie „**Entscheidung**“ bestimmen Ideal-Werte, wie die Ideal-Joystick-Eingaben der MWB, die von der Objektposition ausgehende optimale Rückführungslinie zur Ideallinie oder die Verteilung des Objektregelungsanteils zwischen den MWBn. Diese Entscheidungen dienen anderen Funktionen wiederum als Datenbasis.

In die Unterkategorie „**Informationsvermittlung**“ fallen Funktionen, welche einige dieser Entscheidungen den MWBn präsentieren, um diese mental zu entlasten. Bei diesen Funktionen handelt es sich bspw. um die Anzeige der Ideal-Objektgeschwindigkeit oder die Anzeige der Ideallinie für die Objektregelung. Solche Anzeigen sind aus Rennspielsimulationen

bekannt. Hier zeigen z. B. Linien auf der Strecke dem Spieler die ideale Fahrlinie an. Die ideale Fahrzeuggeschwindigkeit wird dabei bspw. indirekt über eine Einfärbung der Ideallinie oder direkt durch Zahlenwerte vermittelt.

Zudem befinden sich in dieser Unterkategorie Funktionen, welche den MWBn explizit **Empfehlungen** für den Umgang mit Pfadelementen geben. Eine dieser Funktionen empfiehlt den MWBn, die vertikale Objektgeschwindigkeit zu reduzieren oder zu erhöhen bei Kollisionsgefahr mit einem dynamischen Hindernis, in Abhängigkeit von dessen Geschwindigkeit. Eine vergleichbare Funktionalität aus dem Bereich der Fahrzeugführung bieten sogenannte Kreuzungsassistenten. Diese empfehlen dem Fahrer zumeist visuell und akustisch eine Bremsung bei einer drohenden Kollision mit einem Querverkehr.

Ziel der Funktionen dieser Unterkategorie ist es, die Fähigkeiten und Fertigkeiten der MWB nicht zu ersetzen, sondern zu erweitern. Dafür müssen diese Funktionen lernförderlich gestaltet sein, indem diese die MWB beim Erlernen der Tracking-Aufgabe unterstützen, anleiten sowie nützliche Regeln und Navigationshilfen bereit stellen (s. DIN EN ISO 9241-110, 2006; Sarodnick & Brau, 2006). Ebenso soll durch diese Funktionen ein aufgabenangemessenes Tracking-Verhalten durch eine höhere Fehlertoleranz erreicht werden, indem Fehler mit nur minimalem Aufwand schnell korrigiert werden können (s. DIN EN 894-1, 1997; DIN EN ISO 9241-110, 2006). Dafür müssen die Funktionen den MWBn konstruktive Handlungsempfehlungen geben (s. Nielsen, 1994). Durch diese können die MWB schnell reagieren, da eine Interpretation der Situation sowie eine Entscheidung für eine situationsadäquate Reaktion nicht notwendig sind (s. Maier, 2014).

#### **5.2.2.1.1.4 Funktionen der Kategorie „Handlungsausführung“**

Die informationsvermittelnden Funktionen, der zuvor beschriebenen drei Funktionsklassen, wirken indirekt über die MWB auf den MWB-Objekt-Tracking-Umgebung-Regelkreis ein, wohingegen die Funktionen dieser Hauptkategorie diesen direkt beeinflussen. Diese dienen dazu die MWB in genauigkeits- oder geschwindigkeitskritischen Situationen bei der Objektregelung zu unterstützen. Die Entwickler hatten die Automaten unter der einschränkenden Anforderung zu konzipieren, dass die MWB die Objektregelungshoheit behalten sollen und nur in Ausnahmefällen die Objektregelung gänzlich automatisiert werden darf. Daher ist in der Funktionsmatrix keine Funktion aufgelistet, welche beide oder einen MWB dauerhaft aus dem Regelkreis nimmt. Als Ausnahmesituationen sind Notfälle definiert, bei denen die MWB nicht mehr selbst im Stande sind, eine Situation zu deeskalieren. Dies wäre z. B. der Fall, wenn eine unmittelbar bevorstehende Kollision mit einem Hindernis durch die MWB aufgrund deren Reaktionszeit nicht mehr verhindert werden kann. Eine der Funktionen dieser Hauptkategorie

reduziert abrupt in so einem Fall die vertikale Objektgeschwindigkeit soweit bis die MWB selbst die Situation wieder unter Kontrolle haben. Einige Notbremssysteme aus dem Bereich der Fahrzeugführung bieten vergleichbare Funktionalitäten an. Diese führen eine automatisierte Bremsung bis zum Stillstand aus, wenn der Fahrer aufgrund seiner Reaktionszeit und Bremsstärke eine Kollision nicht mehr verhindern kann.

Die Kategorie „Handlungsausführung“ setzt sich aus den Unterkategorien „Veränderung der Joystick-Empfindlichkeit“, „Veränderung des Objektregelungsanteils zwischen den MWBn“ und „Eingriff in die Objektregelung“ zusammen.

Eine der Funktionen der Unterkategorie „*Veränderung der Joystick-Empfindlichkeit*“ agiert sobald das Objekt den Pfad verlässt. Diese verringert die horizontale Empfindlichkeit der Joysticks in Richtung weg vom Pfad, um das Fehlerausmaß so gering wie möglich zu halten.

Die Funktionen der Unterkategorie „*Veränderung des Objektregelungsanteils zwischen den MWBn*“ verändern die Verteilung des prozentualen Anteils der Joystick-Inputs an der Objektregelung zwischen den MWBn, z. B. in Abhängigkeit von ihrer Regelungsgüte.

Die Funktionen der Unterkategorie „*Eingriff in die Objektregelung*“ übernehmen einen prozentualen Anteil oder die vollständige Objektregelung in Notfallsituationen. Beispielsweise regelt eine Funktion das Tracking-Objekt in Abhängigkeit der Objektregelgüte der MWB anteilig bis vollständig wieder in einen sicheren Bereich auf dem Tracking-Pfad, wenn die MWB allein das Verlassen des Pfads nicht mehr abwenden können. Eine vergleichbare Funktionalität ist aus dem Bereich der Fahrzeugführung von Spurhaltesystemen bekannt. Diese führen leichte, für den Fahrer jedoch spürbare Lenkeingriffe in Richtung Spurmitte aus, wenn sich das Fahrzeug ohne gesetzten Blinker einer Spurbegrenzungslinie kritisch annähert.

#### **5.2.2.1.2 Auslegungsmöglichkeiten der Funktionen**

In der Funktionsmatrix sind die Auslegungsanforderungen für jede Funktion der drei Unterkategorien „Informationsvermittlung“ sowie für die der Hauptkategorie „Handlungsausführung“ aufgelistet, die erfüllt werden müssen. Nachfolgend werden verschiedene Auslegungsmöglichkeiten und deren Vorteile beschrieben.

##### **5.2.2.1.2.1 Funktionen zur Informationsvermittlung**

Assistenzsysteme benutzen im Allgemeinen die drei Modalitäten visuell, auditiv und haptisch zur Informationsvermittlung an den Menschen (Hauß & Timpe, 2000; Hedicke, 2000;

Lermer, 2010; Thoma, 2010). Aus technischen Gründen<sup>97</sup> kann im „ATEO Lab System“ keine haptische Informationsvermittlung an die MWB umgesetzt werden, so dass nachfolgend nicht auf diese Modalität eingegangen wird.

Die menschlichen Sinnesmodalitäten sowie kognitiven Kapazitäten bestimmen die Auslegung der einzelnen informationsvermittelnden Funktionen (Vöhringer-Kuhnt, 2008). So existiert nach dem „Multiplen-Ressourcen-Modell“ von Wickens (2002) bei der menschlichen Informationsverarbeitung für jede Modalität (visuell, auditiv), jede Verarbeitungsstufe (Wahrnehmung, Verarbeitung und Handlung) und jeden Verarbeitungstyp (räumlich vs. verbal) eine eigene Kapazität. Danach können Informationen verschiedener Modalitäten, Verarbeitungsstufen und -typen störungsfrei parallel wahrgenommen und verarbeitet werden. Anders verhält es sich mit Informationen einer Modalität, Verarbeitungsstufe oder -typen. Diese stehen im Verarbeitungsprozess in Konkurrenz zueinander. Überschreiten die benötigten Kapazitäten für diese Informationen die vorhandenen Kapazitäten kommt es zu Störungen im Informationsverarbeitungsprozess. Die Tracking-Aufgabe selbst ist eine visuell-manuelle Tätigkeit. Aufgrund des Tracking-Pfadaufbaus und der Dynamik der Tracking-Aufgabe wird vermutlich bereits ein erheblicher Teil der Kapazität für die visuelle Informationsverarbeitung beansprucht. Daher sollten die Funktionen ihre Informationen auch auditiv an die MWB vermitteln, um eine effektive Aufmerksamkeitslenkung und Informationsverarbeitung zu erreichen (s. Campbell et al., 2007; Preim, 1999; Thoma, 2010). Dabei muss beachtet werden, dass die Eignung einer Modalität abhängig von der zu vermittelnden Information ist (Vöhringer-Kuhnt, 2008). Eine Informationsvermittlung birgt die Gefahr, den Menschen von seiner eigentlichen Aufgabe abzulenken (Weller & Schlag, 2002). Dadurch kann eine notwendige Reaktion von diesem zu spät erfolgen, bspw. auf ein plötzlich erscheinendes Hindernis (Vöhringer-Kuhnt, 2010). Deshalb wurde bei den Auslegungen der Funktionen der Funktionsmatrix darauf geachtet, dass diese eine permanente Aufmerksamkeitszuwendung der MWB auf die trackingrelevanten Situationsaspekte gewährleisten.

So ist eine visuelle Informationsvermittlung prädestiniert für kontinuierliche Informationen oder Hinweise an Menschen (Lermer, 2010; Thoma, 2010) mit möglichst geringem Aufmerksamkeitsabzug von der primären Aufgabe (Campbell et al., 2007). Diese besitzt im Vergleich zu einer auditiven Informationsvermittlung eine geringere Signalentdeckungsrate (Campbell et al., 2007). Zusätzlich weist diese einen geringeren Aufforderungscharakter auf,

---

<sup>97</sup> Die verwendeten Joysticks für die MWB in SAM ermöglichen keine haptische Informationsvermittlung. Deshalb hatten die Entwickler in der Untersuchung Leistungsdeterminante „Informationsverfügbarkeit“ (s. Abschnitt 5.2.3) die Aufgabe sich auf visuelle und auditive Informationsvermittlung der Funktionen zu beschränken (s. Anhang G.1.1).

so dass der Mensch diese weitestgehend ignorieren kann (Thoma, 2010). Um die Wahrnehmbarkeit visueller Hinweise zu erhöhen, werden diese teilweise akustisch unterstützt (Lerner, 2010). Dabei handelt es sich häufig um einen abstrakten Warnton (Maier, 2014). Bei der Gestaltung von permanenten oder zumindest häufigen Informationen und Hinweisen sollte darauf geachtet werden, dass diese nicht aufdringlich sind (Thoma, 2010) und damit als störend vom Menschen wahrgenommen werden (Campbell et al., 2007; Lerner et al., 1996). Aufgrund der Eigenschaften einer visuellen Informationsvermittlung ist diese ungeeignet, zuverlässig und schnell die Aufmerksamkeit des Menschen auf sich zu lenken (Lerner, 2010; Maier, 2014). Daher sollten Warnungen bei einer visuellen Hauptaufgabe akustisch erfolgen, wodurch diese mühelos wahrnehmbar sind (Campbell et al., 2007; Scott & Gray, 2008).

Für die *visuelle Informationsvermittlung* existiert eine Vielzahl an Möglichkeiten. Diese kann u. a. in Form von Icons, geometrischen Formen, Texten, Ziffern, Licht oder Einfärbung erfolgen. Durch Erhöhung der Größe, Verwendung bestimmter Farben, wie rot, (Edworthy & Adams, 1996) oder Blinken (Herczeg, 1994, 2009; Lerner, 2010) kann die Salienz (Wickens & McCarley, 2010, zitiert nach Maier, 2014) und damit die Signalentdeckungsrate einer visuellen Informationsvermittlung erhöht werden (Henning, et al., 2008).

Ein *visuelles Icon* ist eine grafische Repräsentation, welches Objekte, Situationen oder Handlungen bildlich vereinfacht darstellt (Campbell, Richmann, Carney, & Lee, 2004; Preim, 1999). Weil ein Icon intuitiv verstanden werden kann (N. Fricke, 2009) und eine Sprachfreiheit besitzt, können sogar komplexere Warninhalte (Lerner, 2010) schnell an den Menschen vermittelt werden (Maier, 2014). Dadurch kann dieser sofort eine situationsadäquate Entscheidung und Handlung ausführen (C. J. Stevens, Brennan, Petocz, & Howell, 2009). Eine Automatik für SAM kann bspw. die MWB auf leicht verständliche Art mittels eines visuellen Icons auf eine vorausliegende Kurve hinweisen und dabei die Richtung sowie Stärke der Kurve darstellen. Icons in dieser Art werden von dem bereits erwähnten Effizienzassistenten im Bereich der Fahrzeugführung verwendet. Ein weiteres Beispiel für ein leicht verständliches Icon aus dem Bereich der Fahrzeugführung ist das Kaffee-Tassen-Icon im Bereich des Kombi-instruments, welches den Fahrer nach längerer Fahrzeit oder erkannter Inaktivität dazu animieren soll, eine Pause zu machen.

Ebenso können ein *visueller Text* oder *visuelle Ziffern* zur Informationsvermittlung verwendet werden. Wobei einzelne Wörter, wie „Achtung“, als unspezifischer Alarm ausgegeben werden können (N. Fricke, 2009). Geeigneter sind kurze Texte, wenn diese eine Information über die Gefahr oder eine Deeskalationsmöglichkeit enthalten. Kurze Texte oder Ziffern können vom Menschen leicht erfasst werden. Diese führen somit bei diesem zu einer

schnellen situationsangemessenen Reaktion. Ist ein Text dem Menschen zudem bekannt, geht die Geschwindigkeit der Informationsverarbeitung in Richtung eines visuellen Icons (Maier, 2014). Eine Automatikfunktion für SAM kann z. B. den MWBn bei einer Kollisionsgefahr mit einem Hindernis die Handlungsempfehlung geben, die vertikale Objektgeschwindigkeit zu verringern, indem diese das Wort „langsamer“ in der Nähe des Tracking-Objekts einblendet. Funktionen dieser Art sind im Bereich der Fahrzeugführung bspw. von Parkassistenten bekannt. Diese zeigen dem Fahrer bei Annäherung an ein Hindernis per Text an „Bitte bremsen“.

Je länger ein Text oder eine Ziffernfolge ist, desto komplexer sind diese und benötigen mehr Verarbeitungszeit (Campbell et al., 2004). In zeitkritischen Situationen ist eine solche Informationsvermittlung daher ungeeignet.

Ferner kann *Licht* oder *Farbe* zur Informationsvermittlung verwendet werden, bspw. in Form eines Punktes, einer Kette oder einer Fläche. Hierdurch kann i. d. R. nur ein unspezifischer Inhalt vermittelt oder ein kritischer Ort hervorgehoben werden (Herczeg, 1994, 2009). Der Mensch muss sich die Bedeutung aus dem Kontext erschließen oder erlernen, was zu einer längeren Verarbeitungszeit führt. Der Vorteil dieser Informationsvermittlung liegt darin, dass diese im Gegensatz zu den anderen visuellen Informationsvermittlungen peripher wahrgenommen und interpretiert werden kann (Maier, 2014). Der Aufmerksamkeitsfokus kann weiterhin bei der primären Aufgabe bleiben (Kienast et al., 2008). Visuelle Informationsvermittlungen deren periphere Wahrnehmung ausreichend ist, verursachen im Vergleich zu auditiven Informationen, wie Tönen, gleiche (Ho & Spence, 2009) oder gar kürzere Reaktionszeiten (Maier, 2014). Eine Automatikfunktion für SAM kann z. B. den Bereich um ein Hindernis farbig hervorheben.<sup>98</sup> Ein Beispiel für diese Art der visuellen Informationsvermittlung aus dem Bereich der Fahrzeugführung ist der Spurwechselassistent, der den Fahrer bei einem durch einen Blinker angezeigten Spurwechselvorhaben vor einem sich im kritischen Bereich der Nachbarspur bewegenden Fahrzeug mittels einer hell leuchtenden, blinkenden LED am Seitenspiegel warnt.

***Auditive Informationsvermittlungen*** entsprechen Auditory Icons, Earcons, einfachen Tönen oder Sprachausgaben. Durch diese kann in zeitkritischen Situationen eine schnelle und zuverlässige Aufmerksamkeitslenkung sowie Reaktion der MWB ausgelöst werden, da diese eine omnidirektionale<sup>99</sup> Wahrnehmbarkeit besitzen (Graham, 1999; Lerner, 2010), welcher sich der Mensch nur schwer entziehen kann (Stanton & Edworthy, 1999; Thoma, 2010).

---

<sup>98</sup> Da die Hindernisse selbst bereits rot sind, kann durch die Farbe kaum mehr Salienz erreicht werden.

<sup>99</sup> positionsunabhängige Wahrnehmbarkeit

Dadurch weisen diese allerdings auch im Vergleich zur visuellen Informationsvermittlung ein höheres Potential auf, störend zu wirken (Lermer, 2010; Marshall, Lee, & Austria, 2007; Thoma, 2010). Außerdem sind diese geeigneter die Kritikalität einer Situation (Belz, Robinson, & Casali, 1999; Scott & Gray, 2008) bzw. die Dringlichkeit einer Handlung zu vermitteln (Marshall et al., 2007). Zudem verursachen auditive Informationsvermittlungen schnellere Reaktionen beim Menschen (Haas & Edworthy, 2006). Bei einer auditiven Informationsvermittlung kann durch Lautstärke oder -höhe die wahrgenommene Dringlichkeit einer Handlung beeinflusst werden (Edworthy & Adams, 1996).

**Auditory Icons** sind gegenständliche Geräusche (Robinson, Lee, & Casali, 2006) mit einer sehr hohen Ähnlichkeit zu bekannten Alltagsgeräuschen (Maier, 2014). Daher weisen diese eine natürliche, im Alltag entstandene Bedeutung auf. Diese Eigenschaft ermöglicht es, den Grund einer Informationsgabe intuitiv verständlich darzubieten (Graham, 1999) und dadurch die kognitiven Kapazitäten des Menschen kaum zu beanspruchen (Belz et al., 1999). Somit sind Auditory Icons besonders geeignet, als Warnung in seltenen und zeitkritischen Situationen eine schnelle adäquate Reaktion beim Menschen zu verursachen (N. Fricke, 2009). Eine Automatikfunktion für SAM kann bspw. bei einer Kollisionsgefahr den Ton einer Hupe oder eines Autounfalls den MWBn zur Warnung darbieten (s. Graham, 1999; McKeown & Isherwood, 2007; C. J. Stevens, Brennan, & Parker, 2004).

Ein **Earcon** stellt eine Abfolge von abstrakten, synthetischen Tönen dar (Lermer, 2010), welche keinen Bezug zu einem Alltagsereignis aufweisen (Brewster, Wright, & Edwards, 1993). Es handelt sich um ein melodisches oder rhythmisches Motiv (Raffaseder, 2002, zitiert nach Maier, 2014). Die Bedeutung muss vom Menschen, aufgrund der vorliegenden Situationsinformationen, zuerst kombiniert oder erlernt werden (Graham, 1999; Lermer, 2010), sonst wirken diese lediglich alarmierend (N. Fricke, 2009). Daher sind Earcons weniger geeignet, in einer hochkritischen Situation als Warnung zu fungieren. Sie sind nicht in der Lage, eine schnelle adäquate Reaktion beim Menschen auszulösen (Campbell et al., 2004).

Ebenso verhält es sich mit **Tönen**, welche aufgrund ihres inhaltlich unspezifischen Charakters ebenfalls anfänglich nur alarmierend wirken (N. Fricke, 2009; Maier, 2014). Der Mensch muss ihre Bedeutung anhand des Kontextes erst erschließen oder zuvor erlernt haben (Graham, 1999), ehe eine situationsadäquate Reaktion erfolgen kann (Maier, 2014). Für komplexe Tonfolgen benötigt er mehr Zeit für die Informationsverarbeitung sowie zum Erlernen der Bedeutung im Vergleich zu einfachen Tönen. Zudem wird ein einzelner Ton eher als Warnung aufgefasst (Campbell et al., 2007). Beides sind Gründe warum bevorzugt einzelne Töne als Warnung verwendet werden (Haas & Edworthy, 2006).

Durch eine bekannte **Sprachausgabe** können Informationen schnell vermittelt werden (Campbell et al. 2004), weil die Bedeutung nicht erst erlernt werden muss (Haas & Edworthy, 2006). Besonders durch einzelne, meist jedoch unspezifische Wörter kann der Mensch alarmiert werden (N. Fricke, 2009). Eine Sprachausgabe wird mit zunehmender Länge immer komplexer, wodurch mehr Verarbeitungszeit benötigt wird (Campbell et al., 2007; Lerner, 2010). Zudem ergibt sich der volle Inhalt erst am Ende der Sprachausgabe (Chang, Lin, Fung, Hwang und Doong, 2008, zitiert nach Maier, 2014). Daher sollte besonders für zeitkritische Situationen nur ein kurzer Text gewählt werden (Campbell et al., 2004), ein Text, der für solche Situationen aufgrund der intuitiven Verständlichkeit geeignet ist. Eine Automatikfunktion für SAM kann z. B. mittels der Sprachausgabe „links halten“ respektive „rechts halten“ die MWB davon abbringen, das Tracking-Objekt in Richtung unterschiedlicher Pfadalternativen zu regeln. Eine vergleichbare Warnung mit Handlungsempfehlung aus dem Bereich der Luftfahrt ist das Bodenannäherungswarnsystem, welches bei einer kritischen Bodenannäherung als Warnmeldung die Sprachausgabe „TERRAIN“ „PULL UP“ ausgibt.

Im Allgemeinen reagiert der Mensch auf Auditory Icons schneller mit einer Handlung als auf Töne (Belz et al., 1999; N. Fricke, 2009; McKeown & Isherwood, 2007) oder kurze Texte einer Sprachausgabe (Graham, 1999). Kurze Texte wiederum verursachen schnellere Reaktionszeiten als Töne (McKeown & Isherwood, 2007). Ebenso sind Auditory Icons geeigneter als Töne zur Vermeidung einer Kollision (N. Fricke, 2009). Töne können zudem am schlechtesten einer Situation zugeordnet werden (McKeown & Isherwood, 2007).

In der Literatur wird vor allem für Warnungen eine **multimodale Informationsvermittlung** gefordert, meist in Form von redundanter Information, um die Signalentdeckungswahrscheinlichkeit zu erhöhen sowie die menschliche Reaktionszeit zu verkürzen (DIN EN 894-1, 1997; Lerner, 2010; Liebig, 1994; Wickens et al., 2004). Multiple Warnungen haben das Potential, durch die Stärken der einzelnen Komponenten deren Schwächen zu kompensieren (N. Fricke, 2009), wie bspw. durch die Kombination der hohen Signalentdeckungswahrscheinlichkeit einer auditiven Warnung mit der Möglichkeit einer längeren Darbietung und schnellen Informationsvermittlung einer visuellen Warnung (Campbell et al., 2007; Thoma, 2010). Die Multimodalität einer Warnung erhöht zudem deren Salienz sowie die durch den Menschen wahrgenommene Kritikalität einer Situation (Maier, 2014). Mit Ausnahme des Auditory Icons verursacht eine multimodale (Auditory Icon oder Ton plus visuelles Icon) im Vergleich zu einer unimodalen Informationsvermittlung (Ton) eine schnellere Reaktionszeit (Belz et al., 1999). Ebenso führt die Kombination aus einem Auditory Icon mit einer LED zu kürzeren Reaktionszeiten als ein Ton plus LED (N. Fricke, 2009). Eine Automatikfunktion für



SAM kann bei einer Kollisionsgefahr mit einem Hindernis neben einem Auditory Icon, wie bspw. in Form des bereits erwähnten Huptons, ein visuelles Icon, z. B. in Form eines Hindernisses, den MWBn zur Warnung darbieten. Im Bereich der Fahrzeugführung existieren Fahrerassistenzsysteme, welche den Fahrer bei drohender Kollision des Fahrzeugs mit einem anderen Verkehrsteilnehmer mittels eines visuellen Icons zusammen mit einem Ton warnen.

#### **5.2.2.1.2.2 Funktionen zur Handlungsausführung**

Für die Auslegungen der Funktionen der Hauptkategorie „Handlungsausführung“ wurden die unter Abschnitt 4.1 beschriebenen Möglichkeiten und deren Eigenschaften berücksichtigt. Daher sollen diese Funktionen nur in aufgabenkritischen Situationen die Objektregelung der MWB für den Zeitraum der Gefahr direkt unterstützen, in Notfallsituationen gänzlich übernehmen (s. Thoma, 2010), um eine schnelle und zuverlässige Deeskalation einer Situation zu bewirken. Weil die MWB die Objektregelungshoheit behalten sollen, soll diesen durch die Funktionen der Unterkategorie „Eingriff in die Objektregelung“ nur so viel Objektregelungsanteil entzogen werden, wie zur Situationsdeeskalation notwendig ist. Dadurch werden die MWB nicht unnötig bevormundet und aus dem Regelkreis herausgenommen.

Generell sollen automatisierte Funktionen die Fähigkeiten, Fertigkeiten, Erfahrungen und Bedürfnisse der einzelnen Menschen in einem MMS berücksichtigen (DIN EN 894-1, 1997; BildschArbV, 1996; Sarodnick & Brau, 2006; Vöhringer-Kuhnt, 2008). Deshalb wird in der Funktionsmatrix eine flexible Veränderung der Joystick-Empfindlichkeit bzw. Zuteilung des Objektregelungsanteils an die MWB respektive an die Funktion im Vergleich zu einer fixen höher gewichtet. Dabei wird eine adaptive Veränderung der Ausgangswerte aufgrund der Regelungsgüte der MWB einer adaptierbaren durch die MWB bevorzugt. Dies zum einen, um die mentale Beanspruchung der MWB nicht zu erhöhen (s. Kaber & Riley, 1999; Kaber et al., 2001). Zum anderen, weil sich der Mensch bei solchen Aufteilungsaufgaben nicht als der bessere Entscheider erwiesen hat (Hilburn et al., 1993; Kaber & Riley, 1999; Scerbo, 1996).

Bei den Funktionen der Unterkategorien „Veränderung der Joystick-Empfindlichkeit“ und „Veränderung des Objektregelungsanteils zwischen den MWBn“ sind die Veränderungen und die Rückstellung zu den Ausgangswerten in der Funktionsmatrix stufenweise vorgesehen. Gleiches gilt für die Funktionen der Unterkategorie „Eingriff in die Objektregelung“ für die Rückgabe des Objektregelungsanteils an die MWB. Durch dieses graduelle Vorgehen sollen sich die MWB an die Veränderungen gewöhnen respektive sich wieder in den Regelkreis zurückorientieren können (Thoma, 2010), um adäquat reagieren zu können (Sarter et al., 1997). Andernfalls ist mit zu starken oder geringen Objektbewegungen durch die MWB zu rechnen, wodurch die Kritikalität einer Situation verstärkt bzw. erneut entstehen kann. Dagegen soll die

Übernahme eines Objektregelungsanteils durch eine Funktion abrupt erfolgen, um nicht unnötig Zeit zu verlieren.

#### 5.2.2.1.2.3 Weitere Auslegungsaspekte

Für einige Unterstützungsaspekte ist es sinnvoll in Form von eigenen Funktionen danach zu unterscheiden, ob diese die Quer- oder Längsregelung des Objekts betreffen. Es ist davon auszugehen, dass die MWB selbst zwischen diesen beiden Objektregelungsachsen differenzieren. Im Bereich der Fahrzeugführung ordnen Fahrer alle Fahraufgaben und Fahrerassistenzsysteme ausnahmslos und eindeutig der Quer- oder Längsregelung eines Fahrzeugs zu (Maier, 2014). Um bei der Konzeptbewertung eine sinnvolle Achsendifferenzierung honorieren zu können, existieren entsprechend für einige allgemeingültige Funktionen zwei achsenbezogene Funktionen. Beispielsweise finden sich in der Funktionsmatrix die drei Feedbackfunktionen zur Güte der allgemeinen, der horizontalen und vertikalen Objektgeschwindigkeit.

Ebenso ist in der Funktionsmatrix bei einigen Funktionen eine individuelle Informationsvermittlung oder Beeinflussung der Joystick-Eingaben vorgesehen. Dadurch können Automatikkonzepte besser bewertet werden, in denen die Fähigkeiten, Fertigkeiten, Erfahrungen und Bedürfnisse der einzelnen MWB berücksichtigt werden. So kann bspw. bei der Pfadverlassensgefahr die Warnung nur an einen MWB gerichtet werden, wenn dieser der Verursacher ist.

Außerdem ist in der Funktionsmatrix für jede Funktion zur Informationsvermittlung sowie der Hauptkategorie „Handlungsausführung“ eine Aktivierung respektive Deaktivierung durch die MWB vorgesehen. Dadurch kann die Konzeption solch einer Funktionseigenschaft ebenfalls honoriert werden. Der Mensch soll technische Unterstützungsmöglichkeiten in einem MMS flexibel nach seinen Fähigkeiten, Fertigkeiten, Bedürfnissen und Präferenzen verwenden können (DIN EN 894-1, 1997; BildschArbV, 1996; Sarodnick & Brau, 2006).

Neben einstufigen existieren mehrstufige Warnkonzepte, sogenannte *Warnkaskaden*, welche zur besseren Situationsbewältigung und positiveren Bewertung durch den Menschen führen (Maier, 2014). In der Funktionsmatrix ist für einige potentiell kritische Pfadelemente ein solcher Kaskadenansatz vorgesehen. So sollen die MWB bspw. einen Hinweis auf ein bevorstehendes Hindernis bereits vor dessen Sichtbarkeit erhalten. Danach kann den MWBn eine Empfehlung zum Umgang mit dem Hindernis vermittelt werden. Geraten die MWB trotzdem in eine Kollisionsgefahr mit diesem Hindernis, sollen die MWB davor gewarnt werden. Reagieren diese darauf nicht adäquat, greift eine Funktion zur Situationsdeeskalation in die Objektregelung direkt ein. Warnkaskaden werden u. a. im Bereich der Fahrzeugführung eingesetzt. So warnen einige Notbremssysteme bei einer Kollisionsgefahr den Fahrer zuerst mittels eines visuellen Icons begleitet von einem Ton. Reagiert der Fahrer nicht adäquat, erfolgt

bei weiterer Abstandsreduktion zum Gefahrenobjekt ein Warnruck. Reagiert der Fahrer weiterhin nicht angemessen reduzieren einige der Funktionen die Geschwindigkeit zuerst anteilig und bei weiterem Bedarf bis zum Stillstand.

Die zuvor sowie im Abschnitt 4.1 beschriebenen Wirkungen der Auslegungsmöglichkeiten von informationsvermittelnden und handlungsausführenden Funktionen weisen keine Allgemeingültigkeit auf, da diese in einem MMS immer durch den Kontext (wie System inklusive anderer automatisierter Funktionen, Benutzer, Aufgaben) beeinflusst werden (s. Contextual Design von Holtzblatt & Hugh, 2014). Bei der Bewertung einzelner Automatikfunktionen muss immer das Gesamtkonzept beachtet werden. Diese ganzheitliche Konzeptbetrachtung bei der Bewertung einzelner Automatikfunktionen ist durch die Aufstellung verschiedener Auslegungsmöglichkeiten der einzelnen Automatikfunktionen und deren Gewichtung sowie in dem Bewertungsfragebogen (s. Abschnitt 5.2.2.2 und Anhang F.3) berücksichtigt. Beispielsweise dient das Item „Output-Modalität“ der Bewertung, ob in einem Konzept der Einfluss verschiedener Informationsmodalitäten auf die Leistung der MWB in einer Situation beachtet wurde.

In der Funktionsmatrix sind für jede Funktion respektive deren Teilaspekte die einzelnen Auslegungsoptionen in Stufen nach deren Eignung für die jeweilige Aufgabe sortiert und gewichtet. Für Funktionen, welche für den gesamten Pfad relevant sind sowie für deren einzelne Funktionsaspekte liegt für die Auslegungsstufe mit der höchsten Eignung ein Gewicht von sechs Punkten vor. Dieses Gewicht leitete sich aus den insgesamt sechs existierenden Pfadelementen ab: 1. Gerade, 2. Kurve, 3. statisches Hindernis 25 %, 4. statisches Hindernis 50 %, 5. dynamisches Hindernis und 6. Gabelung. Für die letzte Stufe "keine Angabe über ..." ist ein Gewicht von einem Punkt vorgesehen, da die Funktion zumindest von den Entwicklern in deren Konzept berücksichtigt wird, wenn auch nicht näher spezifiziert. Zur Bestimmung der Gewichte der dazwischenliegenden Auslegungsstufen bspw. bei weniger als sechs Stufen inklusive der Kategorie "keine Angabe über ..." wurde die maximale Punktzahl durch die Anzahl der Stufen exklusive der Stufe "keine Angabe über ..." dividiert. Das Ergebnis stellt die Punktabstufung zwischen den einzelnen Stufen dar.

Für den Fall, dass für die gesamte Tracking-Aufgabe relevante Funktionen nur für einzelne Pfadelemente in einem Konzept berücksichtigt sind, wurden diese Gewichte je Pfadelement durch sechs dividiert. Somit erhielt die Auslegungsstufe mit der höchsten Eignung ein Gewicht von nur einem Punkt usw.

Bei Funktionen, welche tatsächlich nur für einzelne Pfadelemente relevant sind, wurde wiederum für die höchste Auslegungsstufe ein Gewicht von sechs Punkten vergeben, usw.

(Details zur Bestimmung der Gewichte s. Anhang F.2.4).

Tabelle 22 präsentiert exemplarisch eine informationsvermittelnde Funktion der Funktionsklasse „Informationsanalyse“ mit deren Anforderungen und gewichteten Auslegungsmöglichkeiten. Je eine exemplarische Funktion für die anderen drei Funktionsklassen findet sich im Anhang F.2.5.

*Tabelle 22: Eine exemplarische Funktion der Funktionsklasse „Informationsanalyse“ mit deren Anforderungen und gewichteten Auslegungsoptionen*

Pfadelement	sobald Hindernis/Gabelung im Sichtbereich bis Objekt Hindernis/Gabelung passiert hat		
Situation	Hinderniskollisionsgefahr/Objektregelung auf jeweils andere oder falsche Gabelungspfadalternative, MWB können dies noch abwenden		
Funktions- klasse	Informationsanalyse		
Unterkategorie	Informationsvermittlung		
Funktion	Warnung Gefahr der Kollision/Wahl unterschiedlicher oder falscher Gabelungspfadalternativen		
Anforderung	<p>Es liegt eine unmittelbare Gefahr vor. Die Warnung muss wahrgenommen werden. Eine sehr schnelle Reaktion der MWB ist nötig.</p> <p>Situation kommt nicht häufig vor, somit können MWB nicht überfrachtet werden und auditive Hinweise sind möglich.</p> <p>Die Übermittlung, welche Gefahr droht bzw. was gemacht werden muss, ist nötig.</p> <p>Individuelle Informationsgabe was gemacht werden muss an MWB in Abhängigkeit deren Regelungsgüte sollte möglich sein.</p>		
Komponenten der Funktion	Informationsvermittlung	individuelle Informationsvermittlung bei unterschiedlicher Regelungsgüte	Aktivierung/Deaktivierung durch MWB
Auslegungs- optionen	~ Auditory Icon mit visuellem Icon ~ Auditory Icon mit Lichtquelle/Einfärbung ~ Auditory Icon ~ kurzer Text einer Sprachausgabe Punkt: 6	Punkt: 6	Punkt: 6
	~ visuelles Icon ~ kurzer visueller Text ~ Lichtquelle/Einfärbung ~ Ton mit Lichtquelle/Einfärbung Punkt: 4		
	~ Earcon ~ Ton Punkt: 2		
	~ keine Angabe über Art der Informationsvermittlung Punkt: 1		

Die Funktionsmatrix ist modular aufgebaut. Dadurch können Automatikkonzepte für nur bestimmte Funktionalitäten, Pfadelemente oder Situationen ebenso wie für die Überwachung und/oder Führung von SAM im Allgemeinen bewertet werden.

### 5.2.2.2 Bewertungsfragebogen

Für den „Bewertungsfragebogen für die Güte von Automatikkonzepten“ wurden einige der Items des „Fragebogens zur Bewertung von Assistenzkonzepten“ und der Kategorien des „Prozessmodells zur Evaluation von Assistenz“ von Krinner (2009) übernommen, sprachlich an die bestehende Perspektive auf SAM in der zweiten ATEO-Projektphase angepasst und erweitert. Ebenso wurden neue Items auf Grundlage der in der Literatur, den Normen und den Richtlinien gefundenen Kriterien sowie der Kriterien der Experten von Krinner (2009) generiert. In der Tabelle 23 finden sich die Items des so entstandenen Bewertungsfragebogens. Sie sind als Aussagen formuliert. Die Items umfassen Eigenschaften der Automaten, aber auch der Konzeptdokumentation. Technische Aspekte spielen für die Bewertung der Automatikkonzepte keine Rolle. Den Entwicklern wurde für ihre Konzeptionsaufgabe mitgeteilt, dass die konkrete technische Auslegung keinen Schwerpunkt darstellt.

Im Anhang F.3 finden sich die Items des „Fragebogens zur Bewertung von Assistenzkonzepten“ und des „Prozessmodells zur Evaluation von Assistenz“ von Krinner (2009) mit Darstellung der daraus übernommenen, angepassten, nicht verwendeten oder neu erstellten Items für den „Bewertungsfragebogen der Güte von Automatikkonzepten“. Für jedes Item findet sich eine Begründung untermauert durch die Bewertungskriterien aus der Experten-evaluation von Krinner (2009). Zudem wird die inhaltliche Bedeutung eines Items des „Bewertungsfragebogens Güte von Automatikkonzepten“ durch beispielhaft dargestellte Ausprägungsextreme abgegrenzt. Tabelle 24 stellt das exemplarisch am Item „Anzeigenkomplexität“ dar.

Tabelle 23: Items des „Bewertungsfragebogens für die Güte von Automatikkonzepten“

Items zur Bewertung der Güte von Automatikkonzepten
<b>1. Aspekte der Informationsvermittlung an die MWB</b>
<b>1.1 Informationsanzahl:</b> In dem Konzept ist eine ausgewogene Anzahl an zu vermittelnden visuellen und/oder auditiven Informationen an die MWB vorgesehen.
<b>1.2 Output-Modalität:</b> In dem Konzept ist ein ausgewogenes Verhältnis bei der Vermittlung von visuellen und auditiven Informationen an die MWB vorgesehen. Dabei wird berücksichtigt, dass bereits die Tracking-Aufgabe die visuelle Informationsverarbeitungskapazität beansprucht.
<b>1.3 Aufgabenangemessenheit:</b> In dem Konzept ist eine Vermittlung von aufgabenangemessenen visuellen und/oder auditiven Informationen an die MWB vorgesehen.
<b>1.4 Selbsterklärungs-/Selbstbeschreibungsfähigkeit:</b> In dem Konzept ist eine Vermittlung von verständlichen visuellen und/oder auditiven Informationen an die MWB vorgesehen.
<b>1.5 Anzeigenplatzierung:</b> In dem Konzept ist eine angemessene Anordnung der Anzeigenelemente vorgesehen. Dadurch können die MWB auf die Aufgabenerledigung fokussiert bleiben.
<b>1.6 Anzeigenkomplexität:</b> In dem Konzept ist eine angemessene Komplexität der visuellen Anzeigen vorgesehen.

<b>1.7 Umsetzungseignung (physische Ebene):</b> In dem Konzept sind für SAM angemessene physikalische Umsetzungen der visuellen und/oder auditiven Informationsvermittlungen vorgesehen.
<b>1.8 Instruktion:</b> In dem Konzept ist vorgesehen, den MWBN in der Instruktionsphase auf verständliche Art die relevanten Informationen über die Automaten, deren Funktionen sowie die Bedeutungen von deren Informationsvermittlungen vorzustellen.
<b>2. Aspekt der Empfindung der MWB</b>
<b>2.1 Spaß:</b> Positives emotionales Empfinden bei der Aufgabenausführung wird durch dieses Konzept begünstigt.
<b>3. Aspekte der Dokumentation der Automatikkonzepte</b>
<b>3.1 Namen:</b> Die im Konzept gewählten Namen für die Automaten und deren Funktionen sind für die Kurzbeschreibung der jeweils dahinter liegenden Aufgabe angemessen.
<b>3.2 Begründung:</b> Die im Konzept vorgenommenen Begründungen für die Notwendigkeiten der Automaten sind angemessen beschrieben. Diese stellen die zu lösenden Probleme dar.
<b>3.3 Anforderungen:</b> Die im Konzept aufgestellten Anforderungen an die Automaten und deren Funktionen sind angemessen beschrieben. Diese entsprechen der Formulierung einer Anforderung. Diese spiegeln die zu erfüllenden Eigenschaften der Automaten wider.
<b>3.4 Funktionen:</b> Die im Konzept aufgestellten Funktionen der Automaten sind angemessen beschrieben. Diese spiegeln deren Aufgaben und Wirkmechanismen wider.
<b>3.5 Skizzen</b>
<b>3.5.1 Relevanz:</b> Die Skizzen des Konzepts erläutern Sachverhalte der Automaten oder deren Funktionen, welche über die rein schriftliche Dokumentation nicht vermittelt werden können.
<b>3.5.2 Angemessenheit:</b> Die Skizzen des Konzepts stellen Sachverhalte der Automaten oder deren Funktionen angemessen dar.
<b>3.6 Ablaufschema:</b> Das Konzept beinhaltet eine angemessene Skizzierung des Ablaufs der Automaten und deren Funktionen mit allen relevanten Aspekten.
<b>3.7 Kombinationen/Konflikte:</b> Die im Konzept aufgestellten Kombinationen von Automaten, deren Funktionen oder Komponenten sowie dabei vermutete Konflikte sind angemessen beschrieben.
<b>4. Meta-Eigenschaften der Automatikkonzepte</b>
<b>4.1 Originalität:</b> Das Konzept sieht eine Optimierung der Ausführung der Tracking-Aufgabe durch die MWB vor, ohne diese im Wesentlichen zu verändern.
<b>4.2 Relevanz:</b> Das Konzept beinhaltet keine überflüssigen Aufgaben der Funktionen.
<b>4.3 Redundanz:</b> Das Konzept sieht notwendige Redundanzen von relevanten Aufgaben vor.
<b>4.4 Stimmigkeit:</b>
<b>4.4.1 Konzept:</b> Das Konzept ist in sich widerspruchsfrei.
<b>4.4.2 Kombinationen/Konflikte:</b> Die im Konzept aufgestellten Kombinationen von Automaten, deren Funktionen oder Komponenten sowie dabei vermutete Konflikte sind widerspruchsfrei.

Krinner (2009) verwendete bei dem „Fragebogen zur Bewertung von Assistenzkonzepten“ für die Items eine vierstufige<sup>100</sup> Antwortskala. Diese wurde für den „Bewertungs-

<sup>100</sup> Die von Krinner (2009) verwendete Antwortskala besteht aus folgenden Stufen: 1. trifft völlig zu, 2. trifft eher zu, 3. trifft nicht zu, 4 trifft gar nicht zu.

fragebogen der Güte von Automatikkonzepten“ durch die fünfstufige<sup>101</sup> Antwortskala zur Bewertung von Aussagen von Rohrmann (1978) ersetzt. Hierbei handelt es sich um eine empirisch entwickelte Skala unter Berücksichtigung psycholinguistischer Aspekte mit annähernd äquidistanten Abständen der Stufen. Zudem ist durch die Verwendung von fünf Stufen eine höhere Differenzierung zwischen den Konzepten möglich. Ebenso wurde die zusätzlich verwendete Antwortoption von Krinner (2009) „wird von diesem Konzept nicht beeinflusst“ durch „wird in diesem Konzept nicht thematisiert“ ersetzt. Diese passt besser auf die Inhalte der Bewertungsitems. Jedes Item besitzt somit ein geschlossenes Antwortformat.

Tabelle 24: Item „Anzeigenkomplexität“ des „Bewertungsfragebogens Güte von Automatikkonzepten“ samt Ausprägungsextreme und Begründung

Items	Begründung
<p><b>1.6 Anzeigenkomplexität:</b></p> <p>In dem Konzept ist eine angemessene Komplexität der visuellen Anzeigen vorgesehen.</p> <p><i>(Bei Umsetzung dieses Konzepts würden alle im Konzept vorgesehenen visuellen Anzeigen zusammen auf dem Bildschirm, aber auch jeweils in sich selbst, nicht zu überladen, zu komplex wirken, sondern einfach, sparsam, übersichtlich und effektiv. Wodurch diese wenig ablenkend wären und trotzdem die relevanten Informationen vermitteln könnten.</i></p> <p>vs.</p> <p><i>Alle visuellen Anzeigen würden zusammen auf dem Bildschirm, aber auch jeweils in sich selbst, zu überladen, zu komplex, zu kompliziert wirken, wodurch diese ablenkend wären. Die komplexen oder komplizierten Anzeigen würden zu einer neuen Art der kognitiven Beanspruchung führen.</i></p> <p><i>Oder die visuellen Anzeigen wären zu sparsam, einfach, minimalistisch gehalten, wodurch relevante Informationen nicht damit vermittelt werden könnten.)</i></p>	<p>Dieser Aspekt ist über die Funktionsmatrix nicht ermittelbar.</p> <p>Die Komplexität der Anzeigengestaltung sollte angemessen reduziert sein (DIN EN 894-1, 1997). Dafür sollten die Informationen einfach (DIN EN ISO 15005, 2003), klar und übersichtlich (DIN ISO/IEC 12119, 1995; Sarodnick &amp; Brau, 2006) gehalten sein. Diese sollten nur aus einer geringen Anzahl an Elementen zusammengesetzt sein, welche in sich ein geschlossenes Gebilde ergeben, um leicht wahrgenommen werden zu können (Campbell et al., 2004). Es sollte ein ästhetisches und minimalistisches Design für die Anzeigen verwendet werden (Nielsen, 1994).</p> <p>Zu minimalistisch sollte eine Anzeigengestaltung wiederum nicht sein, denn damit geht ein reduzierter Informationsgehalt einher. Dies vermindert zwar einerseits die kognitive Beanspruchung der MWB, kann andererseits, aufgrund der fehlenden Information, zu einem nicht situationsadäquaten Verhalten des Menschen führen (Thoma, 2010).</p> <p><u>Expertenevaluation von Krinner (2009)</u></p> <p><i>positive Bewertungskriterien für MWB-Assistenzsystemkonzepte:</i></p> <ul style="list-style-type: none"> <li>• „einfache Visualisierung“ (Experte ID-1)</li> <li>• „sparsame, aber u. U. effektive visuelle Hilfen“ (Experte ID-5)</li> <li>• „einfach und effektiv“, was sich auf die Informationsvermittlung bezieht (Experte ID-7)</li> <li>• „einfaches Anzeigensystem mit wenig Ablenkungspotential“ (Experte ID-16)</li> <li>• „überschaubare Anzeige“ (Experte ID-15)</li> <li>• „übersichtliche Anzeigengestaltung“ (Experte ID-18)</li> </ul> <p><i>negative Bewertungskriterien für MWB-Assistenzsystemkonzepte:</i></p> <ul style="list-style-type: none"> <li>• „zu komplex und überladen“, was sich auf die Informationsvermittlung bezieht (Experte ID-7)</li> <li>• „ablenkende, komplizierte Darstellung der Achsenauslenkung“ (Experte ID-1)</li> <li>• „Ablenkung durch komplizierte Hintergründe“ (Experte ID-14)</li> <li>• „zu kompliziert (zusätzliche Pfeile)“ (Experte ID-20)</li> <li>• „Komplexität der Rückmeldung bzw. Visualisierung“ (Experte ID-15)</li> <li>• „Anzeige führt zu neuer Art kognitiver Beanspruchung“ (Experte ID-15)</li> </ul> <p>Dieser Aspekt ist bereits indirekter Bestandteil des „Prozessmodells zur Evaluation von Assistenz“ (Krinner, 2009) durch die Unterkategorie 2.1 „Eigenschaften und Platzierung der Anzeige“.</p>

<sup>101</sup> Die Stufen der Skala von Rohrmann (1978) lauten wie folgt: 1. trifft völlig zu, 2. trifft ziemlich zu, 3. trifft mittelmäßig zu, 4. trifft wenig zu, 5. trifft gar nicht zu.

### **5.2.3 Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“**

Aufbauend auf der Arbeit von Krinner (2009) diente diese Untersuchung der Überprüfung, wie sich die potentielle Leistungsdeterminante „Informationsverfügbarkeit“ auf die Konzeptionstätigkeit von Entwicklern auswirkt. Wie gut Entwickler prospektiv über Automaten mit Systemereignissen umgehen können, hängt von deren Leistungsdeterminanten ab, bspw. die zur Verfügung stehende Information über das System (Wandke & Nachtwei, 2008).

Die potentielle Leistungsdeterminante „Informationsverfügbarkeit“ wurde als uV in Anlehnung an die Leistungsdeterminante „Art der Information über das System“ von Krinner (2009) systematisch variiert und dabei um Informationsquellen zum Operateur von SAM erweitert (s. Abschnitt 5.2.3.2.1.1). Als VPn kamen professionelle Entwickler von komplexen, dynamischen, technischen Systemen zum Einsatz (s. Abschnitt 5.2.3.2.6). Deren Leistung wurde durch die aVn „Quantität“ und „Qualität“, der durch sie konzipierten Automaten für die Überwachung und Führung des SAM-Prozesses, operationalisiert (s. Abschnitt 5.2.3.2.1.2). Zur Erfassung der Werte dieser beiden aVn kam das eigens entwickelte „Bewertungsinstrument der Güte von Automatenkonzepten“ zum Einsatz (s. Abschnitt 5.2.2). Ebenso wurde dafür von der Autorin ein Spezifikationsbogen zur Dokumentation der einzelnen Automaten und deren Funktionen erstellt (s. Abschnitt 5.2.3.2.1.2). Zusätzlich wurden weitere potentiell die Leistung von Entwicklern beeinflussende Variablen als Kontrollvariablen in die Untersuchung mit einbezogen, wie die Kompetenz der Entwickler bei der Entwicklungstätigkeit. Deren Einfluss auf die aV „Qualität“ wurde eliminiert, um den tatsächlichen Einfluss der uV „Informationsverfügbarkeit“ sicherer nachweisen zu können (s. Abschnitt 5.2.3.3.3.2).

#### **5.2.3.1 Fragestellung und Hypothesen**

In diesem Abschnitt soll die Herleitung der fünften Forschungsfragestellung vorgestellt werden, die dieser Untersuchung zugrunde liegt. Hierbei handelt es sich um eine inhaltliche Fragestellung zur Wirkung der potentiellen Leistungsdeterminante „Informationsverfügbarkeit“ auf die Konzeptionsleistung von Entwicklern. Die Aufklärung dieser Fragestellung stellt einen Beitrag zu dem Teilziel „Identifizierung der Leistungsdeterminanten von Entwicklern von komplexen, dynamischen, technischen Systemen“ des ATEO-Projekts dar.

Zusätzlich werden in diesem Abschnitt die aus der Fragestellung abgeleiteten Hypothesen präsentiert. Diese sind als ungerichtete, unspezifische Unterschiedshypothesen sowie eine gerichtete, unspezifische Trendhypothese formuliert.



### 5.2.3.1.1 Herleitung der fünften Forschungsfragestellung

Entwicklungstätigkeiten entsprechen oftmals dem schwersten Problemtyp, der „dialektischen und Synthesebarriere“ (s. Dörner, 1987; Dörner, 1989, Abschnitt 4.2.2.1). Dabei liegen Entwicklern nur unzureichende Informationen über das zu entwickelnde Produkt, den einzelnen Komponenten sowie deren Beziehungen zueinander vor (G. Fricke, 1993; Pietzcker, 2004). Die Methoden und Mittel zur Beeinflussung der einzelnen Komponenten und zur Zielerreichung sind größtenteils unbekannt oder nicht verfügbar (G. Fricke, 1993; Pietzcker, 2004). Die Menge der vorhandenen relevanten Informationen bestimmt die Höhe der Transparenz einer Entwicklungsaufgabe (Krause, 2000). Je weniger relevante Informationen zur Verfügung stehen, desto unsicherer ist das Entwicklungsvorhaben (Hacker & Sachse, 2006). Je mehr relevante Informationen zur Verfügung stehen, desto fundierter ist das Wissen, auf dem die Entwicklung und kritische Bewertung der eigenen Arbeit basiert. Somit hängt u. a. von den vorliegenden Informationen über das zu lösende Problem ab, wie erfolgreich eine Entwicklungstätigkeit sein wird (Kain, 2008). Vor allem bei neuen und situationsabhängigen Aufgaben werden mehr Informationen benötigt (Wübbelmann, 2001). Ebenso beeinflusst die Informationsbedingung die Dokumentation der Lösungen (Ehrlenspiel, 2003).

Bei realen Entwicklungstätigkeiten können Entwickler i. d. R. auf mehrere Informationsquellen zugreifen. Jedoch muss eine Vielfältigkeit an Informationen nicht zu einem umfangreicheren Wissensgewinn führen (Diefenbach, Chien, Lenz, & Hassenzahl, 2013). Die positive Wirkung einer Information auf die Entwicklungsleistung hängt von deren Relevanz und Qualität ab (Wübbelmann, 2001). Wenn zu viele unwichtige Informationen dabei sind, dann können Entwickler auch überfordert werden. Diese können dann nicht mehr zwischen relevanten und irrelevanten Informationen differenzieren. Dabei können unwichtige Aspekte fokussiert werden (Diefenbach et al., 2013). Zugleich können wichtige Informationen in der Menge untergehen, wodurch die Leistung der Entwickler sinkt (Scholl, 1993, zitiert nach Wübbelmann, 2001). Somit können zu wenige, aber auch unangemessene Informationen während der Entwicklungstätigkeit zu einem nicht benutzer- und/oder anwendungsgerechten (Teil-)System führen (Kain, 2008).

In den beiden Modellen „Leistungsdeterminanten von Entwicklern“ mit den Output-Variablen „Zeit“ und „Effizienz“ findet sich die Informationsverfügbarkeit bereits als eine positive Determinante (s. Rocholl, 2010). Somit ergab sich nachfolgende Forschungsfragestellung:

**Forschungsfragestellung 5:**

Welchen Einfluss haben die Art und die Menge der zur Verfügung gestellten Systeminformationen für die Konzeption von Automaten zur Prozessüberwachung und -führung von komplexen, dynamischen, technischen Systemen durch Entwicklerteams auf deren Leistung?

Zur Aufklärung dieser Forschungsfragestellung wurde in der Untersuchung die Wirkung einer unterschiedlichen Anzahl an potentiell relevanten Informationen für die Konzeption von Automaten zur Überwachung und Führung von SAM betrachtet. Der Effekt von zusätzlichen nicht relevanten Informationen wurde noch nicht überprüft. Somit ließen sich für die Forschungsfragestellung 5 nachfolgende Hypothesen formulieren:

**5.1 Forschungshypothesen bezüglich aV „Quantität der Automattikkonzepte“**

Je mehr relevante Informationen aus verschiedenen Perspektiven auf SAM den Entwicklern in der Konzeptionsphase zur Verfügung stehen, desto detailliertere und gewinnbringendere Analysen können diese durchführen. Die Konzeptionsaufgabe sollte mit dem Zuwachs an Wissen transparenter (Krause, 2000) und sicherer (Hacker & Sachse, 2006) werden.

Dadurch sollten die Entwickler besser in der Lage sein, das System zu verstehen und den SAM-Prozess, dessen potentielle Störungen ebenso wie die Wirkungen der potentiellen Lösungen zu antizipieren. Das kann zu einer höheren Anzahl an Anforderungen an die Automaten für SAM führen, welche gleichzeitig präziser formuliert und aufgabenangemessener konzipiert sind. Aus den Anforderungen ergeben sich wiederum die Entwicklungsziele für die Automaten. (Kain, 2008)

Somit sollte eine größere Menge an relevanten Informationen unterschiedlichster Art zu Automattikkonzepten mit einer größeren Anzahl an notwendigen Automaten und Funktionen führen. Andersherum sollten durch eine geringere Informationsmenge weniger relevante Automaten und Funktionen konzipiert werden. Zusätzlich könnten weniger Informationen die Konzeption von nicht relevanten Automaten und Funktionen zur Folge haben (Kain, 2008). Daraus leiteten sich folgende Hypothesen ab:

*F.5.1.1: Die Art und Menge der, den Entwicklern zur Verfügung stehenden, relevanten Informationsquellen für die Konzeptionsphase haben einen Einfluss auf die Anzahl der konzipierten Automaten nach den Spezifikationsbogen<sup>102</sup>.*

*F.5.1.2: Die Art und Menge der, den Entwicklern zur Verfügung stehenden, relevanten Informationsquellen für die Konzeptionsphase haben einen Einfluss auf die Anzahl der konzipierten Funktionen nach den Spezifikationsbogen.*

*F.5.1.3: Die Art und Menge der, den Entwicklern zur Verfügung stehenden, relevanten Informationsquellen für die Konzeptionsphase haben einen Einfluss auf die durchschnittliche Anzahl der konzipierten Funktionen je Automat nach den Spezifikationsbogen.*

*F.5.1.4: Die Art und Menge der, den Entwicklern zur Verfügung stehenden, relevanten Informationsquellen für die Konzeptionsphase haben einen Einfluss auf die Anzahl der konzipierten Funktionen nach der Bewertungsfunktionsmatrix<sup>103</sup>.*

## **5.2 Forschungshypothese bezüglich aV „Qualität der Automatikkonzepte“**

Durch eine größere Menge an relevanten Informationen aus verschiedenen Perspektiven auf SAM und die sich daraus ergebenden umfangreicheren, präziseren sowie angemesseneren Anforderungen sollten die Entwickler zudem aufgabenadäquatere Automaten und deren Funktionen konzipieren. Dabei sollten ebenso potentielle Konflikte zwischen einzelnen Automaten und deren Funktionen besser vorhergesehen und entsprechende Maßnahmen eingeleitet werden. Außerdem sollten die Automaten und deren Funktionen aufgrund des umfangreicheren Wissens der Entwickler detaillierter und angemessener spezifiziert werden. Dadurch sollten die Automatikkonzepte insgesamt eine höhere Güte aufweisen. Daher ergab sich folgende Hypothese:

*F.5.2: Je mehr relevante Informationsquellen unterschiedlichster Art Entwicklern für die Konzeptionsphase zur Verfügung stehen, desto höher fällt die Güte der Automatikkonzepte aus.*

---

<sup>102</sup> Der, von der Autorin erstellte, Spezifikationsbogen diente den Entwickler-VPn in dieser Untersuchung zur Dokumentation der konzipierten Automaten und deren Funktionen (s. Abschnitt 5.2.3.2.1.2).

<sup>103</sup> Die Bewertungsfunktionsmatrix ist eine Komponente des, von der Autorin entwickelten, „Bewertungsinstruments der Güte von Automatikkonzepten“ (s. Abschnitt 5.2.2).

### ***5.3 Kontrollhypothese bezüglich des Einflusses der Kontrollvariablen auf die aV „Qualität der Automatikkonzepte“***

Die Teamleistung von Entwicklern kann ebenfalls durch eine Reihe von anderen Faktoren beeinflusst werden. Einige dieser potentiellen Leistungsdeterminanten sind dem „Modell Leistungsdeterminanten von Entwicklern“ mit der Output-Operationalisierung „Teamleistung“ (s. Abschnitt 5.2.1.1) zu entnehmen. Hierbei handelt es sich bspw. um Input-Variablen, wie die „Gewissenhaftigkeit“ oder das „Leistungsmotiv“ der Entwickler. Ebenso spielen Prozessvariablen eine Rolle, wie die Art der „internen Kommunikation“. Diese Determinanten können als Moderatoren die Beziehung zwischen der Informationsverfügbarkeit und der Teamleistung erhöhen oder abschwächen (Bennett, 2000). Sollten zwischen den Stichproben der zu betrachtenden Informationsbedingungen Unterschiede hinsichtlich dieser Determinanten bestehen, kann sich die Entwicklungsleistung unabhängig von den jeweils verfügbaren Informationen unterscheiden. Aus Platzgründen werden die Kontrollvariablen an dieser Stelle nicht aufgelistet, sondern auf die Erläuterungen und Auflistung im Abschnitt 5.2.3.2.1.3 verwiesen. Die angenommene Wirkung dieser Kontrollvariablen führte zu der nachfolgenden, aus Platzgründen allgemein formulierten, Hypothese:

*K.5.3: Die Kontrollvariablen haben die Beziehung zwischen der uV „Informationsverfügbarkeit“ und der aV „Güte der Automatikkonzepte“ beeinflusst.*

#### **5.2.3.2 Methoden**

##### **5.2.3.2.1 Operationalisierung der Variablen**

Der nachfolgende Abschnitt soll dem Leser dazu dienen, sich einen Überblick über die Operationalisierungen der uV, aVn und der Kontrollvariablen zu verschaffen.

###### **5.2.3.2.1.1 Operationalisierung der unabhängigen Variable**

Die uV dieser Untersuchung stellte die „Informationsverfügbarkeit“ der Entwickler für die Konzeption von Automaten dar. Diese wurde dreifach systematisch variiert. Bei der Ausprägung wurde sich an den von Krinner (2009) und Henkel (2007) verwendeten Informationsquellen orientiert und diese erweitert um Informationsquellen zum Operateur. Die Anzahl der Informationsquellen wurden von Bedingung zu Bedingung erhöht. In der ersten Bedingung erhielten die Entwicklerteams lediglich einen indirekten Systemkontakt, in Form einer schriftlichen Beschreibung der Kernmerkmale von SAM (s. Anhang G.1.2). Dokumente nach Art einer Betriebsanleitung liefern einen Überblick und helfen eine erste Struktur aufzustellen (Röse, 2003).

In der zweiten Bedingung bekamen die Teams neben dieser Beschreibung noch die Möglichkeit direkt das System SAM mit seinem dynamischen Prozess aus verschiedenen Perspektiven kennenzulernen. Dazu diente zum einen eine Videoaufzeichnung der MWB beim kooperativen Tracking vom Start bis zum Ziel. Dadurch sollten die Entwickler einen Eindruck gewinnen, wie MWB mit unterschiedlichen Strategien versuchen die einzelnen Pfadkomponenten zu bewältigen und welche Probleme dabei auftreten. Videos sind besonders geeignet, um eine Interaktion anschaulich zu vermitteln (Diefenbach et al., 2013). Interessante Sequenzen können wiederholt angesehen werden, was bei schnellen, informationsreichen Situationen von Vorteil ist (Sarodnick & Brau, 2006). Neben Videos ermöglicht das eigene Ausführen der Tracking-Aufgabe, das Verhalten der MWB zu antizipieren (s. Goodman, Clarke, Langdon, & Clarkson, 2007). Daher durften die Entwickler selbst das Tracking ausführen. Dies diente dazu, die Perspektive der MWB einzunehmen, um ein noch tieferes Verständnis des dynamischen Prozesses von SAM und seinen Konfliktstellen zu gewinnen. Ferner war es den VPn möglich, den MWBn Fragen zum Tracking und den damit verbundenen Problemen zu stellen. Dafür schlüpfte der VI in die Rolle der MWB. Ebenso wurde dafür der, von Krinner (2009) entwickelte, „MWB-Frage-Antwort-Katalog“ an die neuen SAM-Gegebenheiten angepasst und verwendet, um standardisierte Antworten geben zu können (s. Anhang G.1.2). Durch das Interview können detaillierte Informationen zu spezifischen Fragestellungen (Röse, 2003) und dadurch die Anforderungen der MWB an die Automationen ermittelt werden, was zu einer höheren Systemqualität führen sollte (s. Sarodnick & Brau, 2006). Um Automaten mit einer adäquaten Mensch-Maschine-Interaktion für die MWB zu konzipieren, benötigen die Entwickler die Perspektive der MWB (s. Wiczorek & Zinn, 2008). Diese Bedingung entspricht der Operationalisierung des Faktors „Art der Information über das System“ der ressourcenmaximierten Bedingung der Untersuchung von Krinner (2009) und Henkel (2007).

Die dritte Bedingung bot neben den bisherigen Informationsquellen das Kennenlernen eines Operators von SAM, seines Arbeitsplatzes sowie seiner Aufgaben. Zur Ermittlung der Anforderungen an die Automationen dienen auch die Analysen von Systemen, welche in dem Anwendungsbereich des zu entwickelnden Systems eingesetzt werden (s. Röse, 2003; Sarodnick & Brau, 2006; Wandke, Oed, Metzker, & Nitschke, 2001), wie der OA. Dazu erhielten die Entwickler ein Screenrecording eines exemplarischen OAs, an dem ein Operator SAM überwacht und führt. Des Weiteren sollten die Entwickler selbst als Operateure agieren, um von dessen Tätigkeit und deren Dynamik ein tieferes Verständnis erlangen zu können. Zudem wurde den Entwicklern die Möglichkeit gegeben, den Operator zu interviewen, um vorhandene Fragen aufklären zu können. Hierfür übernahm der VI die Rolle eines Operators.

Zur Standardisierung der Antworten wurde ein „Operator-Frage-Antwort-Katalog“ von der Autorin entwickelt (s. Anhang G.1.2). Dessen weitestgehende Vollständigkeit sowie die des MWB-Katalogs wurden mit einigen Studenten des ATEO-Projekts geprüft. Ebenso wurden in der Untersuchung aufkommende neue Fragen durch die professionellen Entwickler sowie die gegebenen Antworten in den jeweiligen Frage-Antwort-Katalog integriert. Ferner erhielten die Entwickler das Ergebnis der HAA bzgl. der Aufgaben des Operators bei der Überwachung und Führung von SAM (s. Abschnitt 3.3.2 und Anhang C.1.2). Die zu erledigenden Aufgaben, deren Abläufe und Abhängigkeiten sind entscheidend für die Gestaltung eines Systems (Sarodnick & Brau, 2006). Eine Aufgabenanalyse bildet die Grundlage dafür, ob und in welchem Ausmaß eine Aufgabe automatisiert werden sollte (Wiczorek & Zinn, 2008). Ebenso kann daraus abgeleitet werden, wann die MWB eine Information benötigen (s. Sarodnick & Brau, 2006).

Nachfolgend sind die drei Bedingungen überblicksartig dargestellt:

*1. Bedingung (B 1): Systembeschreibung*

*2. Bedingung (B 2): zusätzlich zu 1. verschiedene Informationsquellen zu SAM:*

- *Video von MWBn beim kooperativen Tracking*
- *selbst als MWB agieren*
- *MWB interviewen*

*3. Bedingung (B 3): zusätzlich zu 1. und 2. verschiedene Informationsquellen zum Operator:*

- *Screenrecording der Operateurstätigkeit am OA*
- *selbst als Operator agieren*
- *Operator interviewen*
- *HAA bzgl. der Aufgaben des Operators*

#### **5.2.3.2.1.2 Operationalisierung der abhängigen Variablen**

Um den immer stärker ausgeprägten Konkurrenzkampf der Produkte auf dem Markt bestehen zu können, ist eine gute Leistung der Entwickler äußerst wichtig (Emmanuelides, 1993). Diese Leistung spiegelt sich in der Quantität und Qualität der, durch die Entwickler konzipierten, Automatikkonzepte wider. Unter der Qualität wird die erreichte Güte der konzipierten Automatikkonzepte in Form von Quantität und Qualität der Funktionen sowie deren Dokumentation verstanden. Ermittelt wurde diese mit dem eigens dafür entwickelten „Bewertungsinstrument der Güte von Automatikkonzepten“ (s. Abschnitt 5.2.2). In die Bewertung gingen nur die extra dafür entwickelten Spezifikationsbogen (s. Anhang G.2) der Automaten ein. Pro Automatik sollten die Entwickler einen Bogen erstellen. Darin sollte der Name der Automatik, warum diese nötig ist, welche Anforderungen an diese gestellt werden, die

Spezifikation der einzelnen Funktionen, Skizzen der Informationsvermittlungen an die MWB und Skizzen zur Erleichterung des Konzeptionsverständnisses dokumentiert werden (s. Anhang G.2 und F.3.1, wo im Rahmen der Entstehung des Bewertungsfragebogens die Erfassung dieser Aspekte im Detail begründet wird). Dieser Spezifikationsbogen sollte einen einheitlichen Rahmen für die Konzeptdokumentation bieten, um die Auswertung zu erleichtern. Bei einer Entwicklungstätigkeit wechseln sich die gemeinschaftliche verbale Konzeption von Lösungs-ideen mit deren schriftlichen Beschreibungen und Skizzierung ab (Sachse, Hacker, Leinert, & Riemer, 1999). Die Multimodalität der Lösungsfindung in Form von Sprache und Skizzen stellt eine wesentliche Eigenschaft erfolgreicher Entwicklungstätigkeiten dar (Dylla, 1991).

### **1. aV „Qualität der Automatikkonzepte“:**

- *Güte der Automatikkonzepte*

Die Quantität wird operationalisiert durch die konzipierte Anzahl an Automaten, Anzahl an Funktionen und durchschnittliche Anzahl an Funktionen je Automat nach den Spezifikationsbogen sowie der Anzahl an Funktionen nach der Bewertungsfunktionsmatrix<sup>104</sup>.

### **2. aV „Quantität der Automatikkonzepte“:**

- *Anzahl der Automaten nach den Spezifikationsbogen*
- *Anzahl der Funktionen nach den Spezifikationsbogen*
- *durchschnittliche Anzahl an Funktionen je Automat nach den Spezifikationsbogen*
- *Anzahl der Funktionen nach der Bewertungsfunktionsmatrix*

#### **5.2.3.2.1.3 Operationalisierung der Kontrollvariablen**

Durch die Untersuchung sollte ermittelt werden, unter welcher Ausprägung der uV die besten Automatikkonzepte erstellt werden. Zur Analyse des Einflusses der uV auf das Konzeptionsergebnis wurde eine Vielzahl von Kontrollvariablen in Form von Input- und Prozessvariablen erfasst. Hierbei handelt es sich um Variablen, welche in der Literatur bereits als potentielle Leistungsdeterminanten von Entwicklern identifiziert oder zumindest theoretisch als solche argumentiert wurden (s. Acuña, Gómez, & Juristo, 2009; Barrick & Mount, 1991; Barrick, Stewart, Neubert, & Mount, 1998; Bell, 2007; G. Fricke, 1993; Koltko-Rivera, 2004; Peeters et al., 2008; Peeters, Tuijl, Rutte, & Reymen, 2006; Poling, Woehr, Arciniega, &

---

<sup>104</sup> Die Bewertungsfunktionsmatrix ist eine Komponente des, von der Autorin entwickelten, „Bewertungsinstruments der Güte von Automatikkonzepten“ (s. Abschnitt 5.2.2).

Gorman, 2006; Sfetsos, Stamelos, Angelis, & Deligiannis, 2009; Sonnentag, 1994). Diese Kontrollvariablen sollten dazu dienen post hoc einen Teil der Varianz der  $\alpha_V$  „Güte der Automatikkonzepte“ aufzuklären (s. Jo, 2002), um die gefundenen Ergebnisse eindeutiger interpretieren und somit sicherere Aussagen über den Effekt der  $u_V$  treffen zu können (Huber et al., 2008; Kain & Nachtwei, 2009; C. Meyer et al., 2009; Nachtwei & Kain, 2010). Zugleich wurden diese Variablen verwendet, um das „Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme“ mit der Output-Variable „Teamleistung“ (s. Abschnitt 5.2.1) zu validieren und zu erweitern, um so ein empirisch validiertes Modell speziell für Entwickler von Automationen zu gewinnen. Für eine erste Validierung wurden aus diesem Modell die Input-Faktoren „Gewissenhaftigkeit“ und „Leistungsmotiv“ sowie der Prozessfaktor „interne Kommunikation“ als potentielle Leistungsdeterminanten ausgewählt.

Gewissenhaftigkeit ist einer der fünf allgemein hin anerkannten (Rammstedt, Koch, Borg, & Reitz, 2004) Hauptfaktoren der Persönlichkeit (Costa & McCrae, 1997; McCrae & Costa, 1987), der sogenannten Big Five (Goldberg, 1981). Gewissenhaftigkeit entspricht u. a. dem Streben nach Kompetenz und Leistung (Rammstedt & John, 2007). Personen mit einer hohen Gewissenhaftigkeit sind zielorientiert (Peeters et al., 2006). Diese Persönlichkeitsdimension ist über die Zeit und Situation sehr stabil (McCrae & Costa, 1990).

Das Leistungsmotiv ist das Bestreben, ein herausforderndes Ziel zu erreichen (Byrne, et al., 2004). Dieses wird in eine Annäherungs- und eine Vermeidungstendenz untergliedert (Atkinson, 1957). Mit der Tendenz zur Annäherung ist das Streben nach Erfolg gemeint, wenn eine Hoffnung auf Erfolg besteht (Lang & Fries, 2006). Wohingegen die Vermeidungstendenz sich auf einen möglichen Misserfolg bezieht, da eine Furcht vor Misserfolg vorliegt (Lang & Fries, 2006). Beide Tendenzen motivieren menschliches Verhalten und kommen in unterschiedlichen Ausprägungskombinationen vor. Es handelt sich um Dispositionen, d. h. eine zeitlich überdauernde Neigung, bestimmtes Verhalten in bestimmten Situationen zu zeigen (Huber et al., 2008). Personen mit einer hoch ausgeprägten Annäherungstendenz an Erfolg vollbringen eine bessere Leistung (Spangler, 1992, zitiert nach Lang & Fries, 2006). Aufgrund von situativen Bedingungen wird ein Leistungsmotiv zu einer Leistungsmotivation angeregt. Diese schlägt sich in einem zweckorientierten Handeln nieder (Müsseler & Prinz, 2002). Je höher die Motivation ist, desto besser fällt die Leistung aus (Blau, 1993). Daher wurde zudem die Leistungsmotivation der Entwickler über deren subjektive Wichtigkeit, eine gute Teamleistung zu erreichen, als Kontrollvariable erfasst.



Die interne Kommunikation ist ein facettenreicher Faktor, welcher alle Interaktionen in einem Entwicklerteam umfasst, wie aufgabenbezogene und sozio-emotionale Kommunikation (Badke-Schaub & Stempfle, 2003). Deshalb wurden u. a.<sup>105</sup> die aufgabenbezogenen inhaltlichen Aussagen der Entwickler als Kontrollvariable erhoben. Zudem wurde deren Explorationstiefe ermittelt, da diese einen positiven Einfluss auf die Teamleistung haben soll (van Knippenberg & Schippers, 2007). Ebenso wurde die Wirkung der Positivität der sozio-emotionalen Aussage auf die aV kontrolliert. Dabei wurde explizit der Zusammenhang zwischen der Anzahl an eingesetzten Witzen durch die Entwickler bei der Teamkommunikation mit deren Leistung betrachtet. Humor ist ein geeignetes Mittel, um in einer sozialen Interaktion eine gegensätzliche Meinung zu vermitteln, ohne angreifend zu wirken, Reaktanz beim Gegenüber oder gar Konflikte zu verursachen (G. Schwarz, 2008). Ebenso kann bereits entstandene Spannung durch Humor zum Erhalt einer positiven Beziehung wieder abgebaut werden (Becker-Beck, 1994). Somit kann der Einsatz von Witzen auch als Instrument des Konfliktmanagements gesehen werden, welches ebenfalls im positiven Zusammenhang mit dem Entwicklungserfolg steht (s. "Gesamtmodell der Leistungsdeterminanten von Entwicklern" Anhang E; Pinto & Prescott, 1988).

Zur Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“ mit der Output-Variable „Teamleistung“ wurden aus dem Gesamtmodell die Faktoren „Arbeitserfahrung“ und „-kompetenz“ als Kontrollvariablen herangezogen. Dafür wurde die subjektive Erfahrung der Entwickler mit komplexen, dynamischen Systemen, insbesondere der Automationen und der Funktionsteilung ermittelt. Zusätzlich wurde der Anteil erfasst, den die Problemidentifizierung bis hin zur Konzeption in ihrem Arbeitsalltag ausmacht. Die Arbeitskompetenz entspricht dem Wissen und der Fähigkeit, das Vorgehen bei einer Entwicklungstätigkeit zu planen, zu steuern, zu kontrollieren und dabei das zu bewältigende Problem sowie Lösungsalternativen zu erkennen (s. G. Fricke, 1993; Högl, 1998). Diese wurde über die subjektive Kompetenz der Entwickler bei der Entwicklungstätigkeit erhoben.

Zusätzlich wurde als Kontrollvariable die Vertrautheit im Team verwendet. In vertrauten Teams läuft die Kommunikation zielorientierter und effizienter, da keine Zeit zum gegenseitigen Kennenlernen benötigt wird (Harrison, Mohammed, McGrath, Florey, &

---

<sup>105</sup> Weitere Variablen zur ausdifferenzierten Operationalisierung der internen Kommunikation, wie steuerungsbezogene Aussagen der Entwickler, ebenso wie die weiteren vier Persönlichkeitsdimensionen „Neurotizismus“, „Extraversion“, „Offenheit für Erfahrungen“ und „Verträglichkeit“ wurden erhoben. Deren Analysen erfolgten im Rahmen der Diplomarbeit von Dodt (2012), angeleitet durch die Autorin. Da sich bei diesen Variablen keine signifikanten und praktisch bedeutsamen Zusammenhänge zur aV „Güte der Automatikkonzepte“ nachweisen ließen, werden diese aus Platzgründen in dieser Arbeit nicht erwähnt. Die Variablen, bei denen sich tendenziell signifikante und dabei praktisch bedeutsame Zusammenhänge mit der aV zeigen, werden im Anhang G.6.5 präsentiert.

Vanderstoep, 2003). Dadurch kann eine höhere Gruppenleistung erreicht werden (Reagans, Argote, & Brooks, 2005).

Tabelle 25: Die, für die Datenanalyse, relevanten erhobenen Variablen respektive eingesetzten Fragebogen, die verwendeten Antwortskalen und Messzeitpunkte in der Untersuchung „Informationsverfügbarkeit“

Variable (Fragebogen)	Antwortskala	Prä	Vi	Post	On
<b>Input-Variablen</b> (individuell erhoben und pro Team zum M aggregiert)					
<b>Team</b>					
Alter	offen			X	X
Wichtigkeit der Teamleistung	VAS (-50 bis 50)			X	
Gewissenhaftigkeit (BFI-45 von Rammstedt, 1997)	VL (5-stufig)				X
Leistungsmotiv (AMS-R von Lang & Fries, 2006)					
<i>Leistungsmotiv Hoffnung auf Erfolg (HE)</i>	VL (4-stufig)				X
<i>Leistungsmotiv Furcht vor Misserfolg (FM)</i>	VL (4-stufig)				X
Kompetenz bei der Entwicklungstätigkeit	VL (5-stufig)			X	
Erfahrungsindex (Werte aus nachfolgenden fünf Items zum M aggregiert)					
<i>Interesse Funktionsteilung</i>	VL (5-stufig)			X	
<i>Interesse Automatisierung</i>	VL (5-stufig)			X	
<i>Erfahrung dynamische Systeme</i>	VL (5-stufig)				X
<i>Erfahrung komplexe Systeme</i>	VL (5-stufig)				X
<i>Erfahrung komplexe, dynamische Systeme</i>	VL (5-stufig)				X
Vertrautheitsindex (Werte aus nachfolgenden vier Items zum M aggregiert)					
<i>Bekanntheit Teammitglied niedrigere Nr.</i>	VL (4-stufig)			X	
<i>bisherige Zusammenarbeit Teammitglied niedrigere Nr.</i>	VL (5-stufig)			X	
<i>Bekanntheit Teammitglied höhere Nr.</i>	VL (4-stufig)			X	
<i>bisherige Zusammenarbeit Teammitglied höhere Nr.</i>	VL (5-stufig)			X	
prozentualer Anteil Problemidentifizierung bis Konzeption im Berufsalltag (Werte aus nachfolgenden vier Items zum S aggregiert)					
<i>Problemidentifizierung</i>	%			X	
<i>Lösungssuche</i>	%			X	
<i>Auswahl der geeignetsten Lösungsidee/n</i>	%			X	
<i>Konzeption</i>	%			X	
<b>Aufgabe</b>					
mentale Beanspruchung (RSME <sub>A</sub> )	VAS (0 bis 150)			X	
Aufgabenkomplexität	VAS (-50 bis 50)			X	
Informationsquellen ausreichend	VL (5-stufig)			X	
Höhe Wichtigkeit je Informationsquelle	VL (5-stufig)			X	
Rangreihe Wichtigkeit der Informationsquellen	R			X	

Variable (Fragebogen)	Antwortskala	Prä	Vi	Post	On
<b>Prozessvariablen</b> (individuell erhoben und pro Team zum M aggregiert)					
Stimmung	Kunin (7-stufig)	X		X	
<b>Variablen des Instruments zur Kodierung von Diskussionen</b> (kurz IKD; Schermuly, 2007, 2010; Schermuly & Scholl, 2010; Schermuly, Schröder, Nachtwei, & Scholl, 2010; pro Team als S erhoben)					
Variablen der funktionalen Bedeutung einer Aussage:					
<i>aufgabenbezogene inhaltliche Aussage</i>	H		X		
<i>Positivität der sozio-emotionalen Aussagen</i> (aggregiert aus Anzahl sozio-emotional positiver minus negativer Aussagen)					
<i>sozio-emotional positive Aussage</i>	H		X		
<i>sozio-emotional negative Aussage</i>	H		X		
<i>Witz</i> (Erweiterung IKD)	H		X		
Variable der Nebenkategorien einer Aussage:					
<i>Elaborationstiefe</i> (Erweiterung IKD durch Meyer & Schermuly [2011] nach Homan [2007])	VL (3-stufig)		X		
Anmerkungen. Prä = Vorbefragung; Vi = Videoaufnahme der Konzeption; Post = Nachbefragung; On = Online-Befragung; VAS = visuelle Analogskala; VL = Ratingskala mit verbalen Labels für jede Stufe; H = Häufigkeit; R = Rangreihe; % = Prozentangaben; M = Mittelwert; S = Summe					

Außerdem wurde das Alter der Entwickler erhoben und geprüft, ob eventuell vorhandene Unterschiede zwischen den Stichproben bzgl. des durchschnittlichen Teamalters einen Einfluss auf die aV „Güte der Automatikkonzepte“ haben.

Ferner wurde ein Manipulation-Check durchgeführt. Es wurde geprüft, ob die, den Entwicklern zur Verfügung gestellten, Informationsquellen tatsächlich als relevant für die Konzeption von Automaten für SAM empfunden wurden. Dafür sollten die Entwickler deren Wichtigkeit einschätzen und diesbzgl. in eine Rangreihe bringen. Dabei wurde auch gefragt, ob die Informationsquellen ausreichend zur Bewältigung der Aufgabe waren.

Darüber hinaus wurde überprüft, ob die Konzeptionsaufgabe herausfordernd für professionelle Entwickler ist. Dafür mussten die Entwickler die eigene Beanspruchung durch die Aufgabe sowie deren Komplexität einschätzen. Eine beanspruchende Aufgabe sollte zu einer negativen Veränderung der Stimmung führen (Kaufmann et al., 1982). Deshalb wurde die subjektive Stimmung der Entwickler vor und nach der Konzeptionstätigkeit erhoben.

Bei allen Kontrollvariablen wurde darauf geachtet, dass diese ökonomisch messbar sind und trotzdem valide Aussagen über die Variablen selbst erlauben. Die inhaltlichen und sozio-emotionalen Prozessvariablen wurden anhand der Videoaufzeichnungen der Entwicklerteams bei der Lösung der Konzeptionsaufgabe ermittelt. Alle anderen Kontrollvariablen wurden über Items erhoben, welche die drei Entwickler unabhängig voneinander beantworteten. Die Items

weisen unterschiedliche Antwortskalen auf. In Tabelle 25 sind die im Rahmen dieser Arbeit relevanten Kontrollvariablen der Untersuchung samt Erhebungsinstrument, Antwortskala und Messzeitpunkt aufgelistet. Anhang G.4 bietet eine Übersicht über alle Kontrollvariablen der Untersuchung. Die konkreten Formulierungen der Items und deren Antwortskalen sind im Anhang G.5 nachzulesen. Für die Online-Befragung wurde das Online-Umfrage Tool Unipark verwendet.

### **5.2.3.2.2 Erhebungsinstrumente**

#### **5.2.3.2.2.1 Big Five Inventory, deutsche 45-Item Version**

Die etablierte deutsche 45-Item Version des „Big Five Inventory“ (BFI-45) von Rammstedt (1997) diente der Erhebung der Hauptpersönlichkeitsdimension „Gewissenhaftigkeit“. Dieses Instrument weist adäquate psychometrische Eigenschaften für die einzelnen Dimensionen auf. (Details s. Anhang G.3.2)

#### **5.2.3.2.2.2 Achievement Motive Scale-Revised**

Mittels der „Achievement Motive Scale-Revised“ (AMS-R) von Lang und Fries (2006) wurde die Leistungsmotivation der MWB erhoben. Dieses Instrument erfasst die Dimensionen „Hoffnung auf Erfolg“ (HE) und „Furcht vor Misserfolg“ (FM). Diese Kurzsкала weist adäquate psychometrische Eigenschaften auf. (Details s. Anhang A.2.1)

#### **5.2.3.2.2.3 Instrument zur Kodierung von Diskussionen**

Für die Quantifizierung des Großteils der Prozessvariablen wurde das „Instrument zur Kodierung von Diskussionen“ (IKD; Schermuly, 2007, 2010; Schermuly & Scholl, 2010; Schermuly et al., 2010) gewählt (Details s. Anhang G.3.3). Das IKD ermöglicht, u. a. die funktionale Bedeutung der Entwickleraussagen der inhaltlichen, steuerungsbezogenen und sozio-emotionalen Ebene zuzuordnen. Zudem wurde das Instrument durch die Autorin um die Erhebung der Variable „Witzanzahl“ erweitert.

#### **5.2.3.2.3 Aufbau der Untersuchung**

Die Untersuchung wurde quer durch Deutschland an den jeweiligen Arbeitsstätten der Entwickler durchgeführt. Bei den Untersuchungsräumen handelte es sich um reizarme Meetingräume, welche eine ungestörte Untersuchung und einen nahezu gleichen Untersuchungsaufbau (s. Abbildung 23) gewährleisteten. Der V1 setzte sich während der Untersuchung so weit wie möglich von den VPn weg.

In der **Bedingung 1** wurde ein VL-Laptop (13“ DELL XPS M1333, Betriebssystem: Microsoft Windows 7 Professional 32bit, Bildschirmauflösung: 1280 x 800 mit 60 Hz) sowie ein daran angeschlossener Beamer (ACER DLP Projector) für die Präsentation der Instruktionen per Microsoft Office PowerPoint 2007 verwendet. Der VL-Laptop wurde vor dem VI platziert. Als Eingabeinstrument diente dem VI eine Maus.

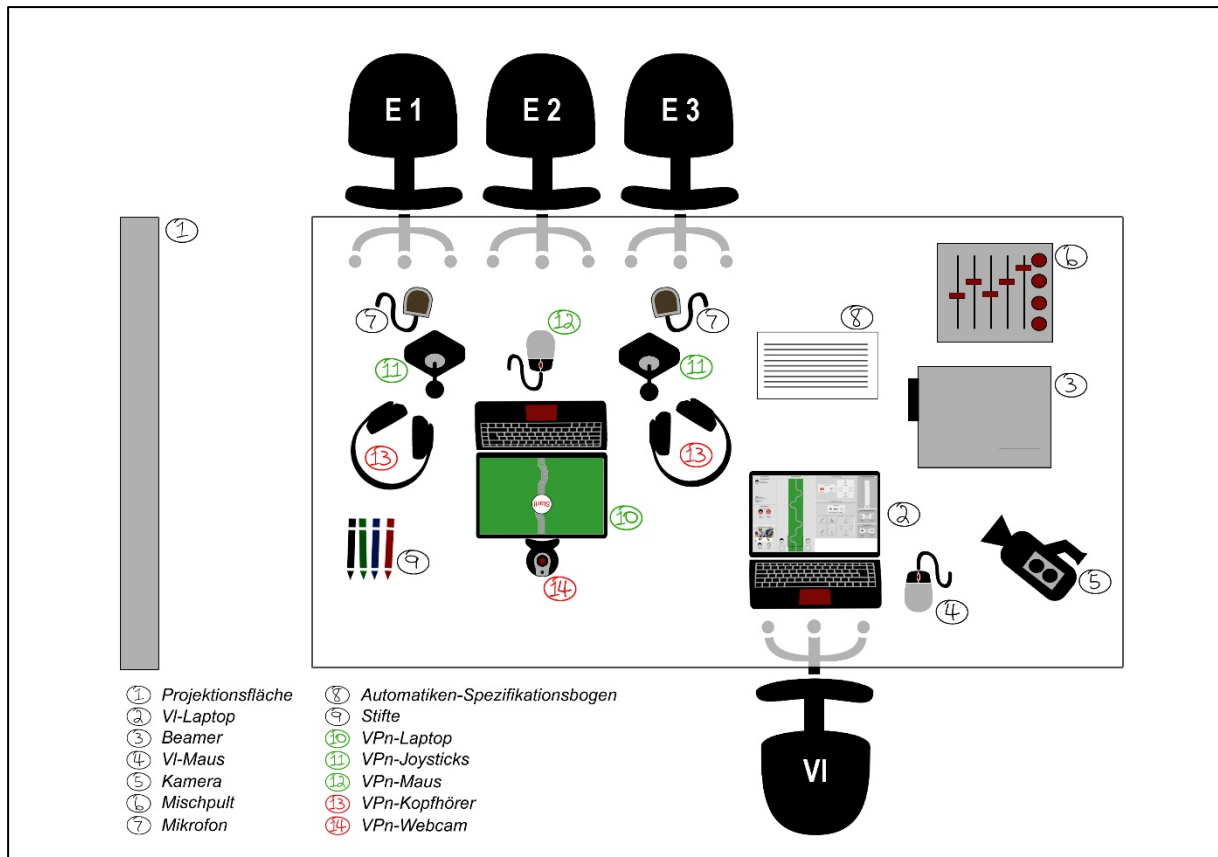


Abbildung 23: Schematische Zeichnung eines exemplarischen Untersuchungsaufbaus

Die aktive Phase der Entwickler wurde per Video mit einer Kamera (Sony DCR-TRV15E PAL mit Digital Zoom 120x) aufgenommen. Um nicht die Arbeit der VPn zu stark durch die Kamera zu beeinträchtigen und dadurch Verhaltensänderungen zu bewirken, wurde diese maximal weit weg und unauffällig aufgestellt. Diese Kamera wurde an den VL-Laptop per FireWirekabel angeschlossen. Von dem Laptop wurden die Kameraaufnahmen über Windows Movie Maker im Format Windows Media Video importiert und mit Beendigung der Aufnahme automatisch gespeichert. Zudem wurden an die Kamera über ein Mischpult (Compact 4 Soundcraft) mittels Audioadapterkabel zwei Grenzflächenmikrofone (1x Sennheiser E912 & 1x Mc Crypt CM-550) angeschlossen, um die Qualität der Sprachaufnahme zu erhöhen. Die beiden Mikrofone wurden so vor den VPn platziert, dass deren sprachliche Äußerungen maximal verständlich aufgezeichnet werden konnten.

In der **Bedingung 2** wurde zusätzlich zu diesem Aufbau ein VPn-Laptop (13“ Sony VAIO PCG-5T1M, Betriebssystem: Microsoft Windows 7 Professional 32bit, Bildschirm-auflösung: 1280 x 800 mit 60 Hz) verwendet, zum Agieren der VPn als MWB. Dieser wurde vor den VPn positioniert. An diesem wurden zwei Joysticks (Logitech Attack 3 Joystick) und eine Maus angeschlossen.

In der **Bedingung 3** wurden, über diesen Aufbau hinaus, die beiden Laptops miteinander per Crossoverkabel verbunden, damit die VPn als Operateure agieren konnten. Mittels des VI-Laptops konnten die VPn die Tätigkeit des Operateurs übernehmen. Der OA wurde den VPn über die Beamer-Projektion präsentiert. Die an dem VI-Laptop angeschlossene Maus wurde den VPn zur Bedienung des OAs zur Verfügung gestellt. An dem VPn-Laptop konnten die VPn weiterhin als MWB fungieren. Damit diese dabei auditive Informationen vom Operateur erhalten konnten, wurden an diesem Laptop zwei Kopfhörer (Philips SHP1900) angeschlossen. Damit im OA die Videodaten der beiden als MWB fungierenden Entwickler dargestellt werden konnten, wurde eine Webkamera (Logitech Pro 9000 QuickCam mit 2-Megapixel und Autofokus) an den VI-Laptop angeschlossen und mittig über dem VPn-Laptopmonitor positioniert.

#### 5.2.3.2.4 Experimentelles Design

In der Untersuchung wurde der Effekt einer experimentellen systematischen Manipulation der potentiellen Leistungsdeterminante „Informationsverfügbarkeit“ von Entwicklern auf die Konzeption von Automaten für die Prozessüberwachung und -führung von SAM 2.0 überprüft. Das experimentelle Design war ein einfaktorieller Versuchsplan. Bei der uV „Informationsverfügbarkeit“ handelte es sich um einen dreifach gestuften Gruppierungsfaktor. Somit geschah der Nachweis über den Effekt mittels eines Between-Groups-Designs.

Die Zuordnung eines Entwicklers zu einem Team erfolgte nicht zufällig, da jedes beteiligte Unternehmen respektive Lehrstuhl einer Universität ein bis mehrere Teams à drei Personen stellte. Die Teammitglieder mussten sich jedoch nicht kennen. Des Weiteren fand die Zuordnung zu einer der drei Bedingungen aufgrund von anfänglichen technischen Schwierigkeiten<sup>106</sup> bzgl. des Versuchsmaterials der Bedingungen 2 und 3 nicht gänzlich zufällig statt.

---

<sup>106</sup> Insbesondere die fehlerfreie Fertigstellung der mobilen Variante des OAs zur Selbstaübung der Operateurstätigkeiten durch die Entwickler für die Bedingung 3 verzögerte sich, so dass diese Bedingung erst in den letzteren Untersuchungsdurchgängen durchlaufen werden konnte. Durch Vergabe der Implementierungsaufgaben an Studenten der Informatik ging auf Seiten der Operateursperspektive die Anforderung verloren, dass der Operateursarbeitsplatz flexibel auf verschiedenen Rechnern und Bildschirmgrößen eingesetzt werden können soll. Der von Nachtwei (2011a) konzipierte und durch Leonard (2010, 2013) implementierte Operateursarbeitsplatz war nur auf einem 30“ Monitor anwendbar. Deshalb musste wenige Monate vor Untersuchungsbeginn eine weitere Studienarbeit an einen Studenten, Hasselmann (2010) vergeben werden, welcher eine mobile Version von diesem Operateursarbeitsplatz implementierte (s. Abschnitt 3.3.2.1).

Aufgrund des Designs handelte es sich um eine Laborstudie, durchgeführt im täglichen Arbeitsfeld der Entwickler. Durch diese, für die Entwickler gewohnte, Atmosphäre sollte der Bezug zu realen Entwicklungstätigkeiten erhöht werden. Ebenso sollte dadurch die Wirkungsüberprüfung einiger Kontrollvariablen zu sichereren Ergebnissen führen. Untersuchungen haben gezeigt, dass manche individuelle Variablen, wie die Big Five, aber auch deren Kombination<sup>107</sup> im Team in Feldstudien einen höheren und zumeist erst bedeutsamen Zusammenhang aufweisen als im Labor. So zeigt sich bspw. bei der durchschnittlichen Ausprägung der Team-Verträglichkeit nur in Felduntersuchungen ein positiver Zusammenhang zur Teamleistung (Bell, 2004, 2007).

#### **5.2.3.2.5 Durchführung der Untersuchung**

Die Datenerhebung fand in Einzelsitzungen mit je einem Entwicklerteam statt, selbst wenn ein Unternehmen mehrere Teams stellte. Gründe dafür waren technische Restriktionen. Zudem sollte ein Entwicklerteam die Untersuchung in einer ungestörten Atmosphäre ohne Zeitdruck absolvieren.

##### **5.2.3.2.5.1 Probendurchläufe**

Vor Beginn der Untersuchung wurden vier Pilotuntersuchungen mit je einem Drei-Personen-Team durchgeführt. Die Erste und Letzte fand mit Studenten am Institut für Psychologie des Lehrstuhls Ingenieurpsychologie/ Kognitive Ergonomie der Humboldt-Universität zu Berlin statt. Die anderen beiden wurden in Hamburg bei der Airbus Deutschland GmbH mit professionellen Entwicklern durchgeführt. Die Probendurchläufe dienten dazu, den geplanten Ablauf und die Materialien hinsichtlich Funktionsfähigkeit respektive Verbesserungsbedarf zu überprüfen. Bei der Gestaltung des Untersuchungsablaufs und der Materialien wurde sich an den Untersuchungen von Krinner (2009) orientiert. Es zeigten sich einige Schwachstellen, welche von Pilotuntersuchung zu Pilotuntersuchung angepasst wurden. So musste bspw. aus Zeitgründen auf die Verwendung eines Grafiktablets zur Anfertigung von Skizzen, den Informationsquellen „Logfile der Aktionen der MWB“ und „Logfile der Aktionen des Operators“ verzichtet und einige der Fragen per Online-Nachbefragung erhoben werden. Gleichzeitig trainierte die Autorin die Rolle des Vls sowie den Ablauf ein, um in der Untersuchung selbst keinen Störeinfluss durch unterschiedliches Verhalten bei den einzelnen Teams

---

<sup>107</sup> Kombinationen werden operationalisiert in Form von Höhe (Mittelwert, Summe), Heterogenität (Varianz, Standardabweichung) oder den Extremen (Minimum, Maximum) der Werte der Teammitglieder (Bell, 2007; Peeters et al., 2008; Peeters et al., 2006; Poling et al., 2005).

zu verursachen. Der Ablauf entsprach dem von Bedingung 3 ohne die Informationsquellen „selbst als MWB agieren“ oder „selbst als Operateur agieren“.

#### **5.2.3.2.5.2 Ablauf der Untersuchung**

Die Untersuchung wurde quer durch Deutschland an den jeweiligen Arbeitsstätten der Entwickler durch die Autorin durchgeführt. Je nach Bedingung und gestellten Fragen der VPn dauerte ein Untersuchungsdurchlauf zwischen dreieinhalb und sechs Stunden exklusive einer Stunde Vorbereitungszeit. Ein Untersuchungsablauf lässt sich in fünf Phasen einteilen: Einführung der Entwickler, Exploration der Informationsquellen, Konzeption der Automaten, Abschlussbefragung sowie Online-Befragung. Im Folgenden wird der chronologische Ablauf der Untersuchung vorgestellt, welcher zur Gewinnung eines Überblicks in Abbildung 24 dargestellt ist.

##### ***Phase 1: Einführung der Entwickler***

Die Entwickler kamen mit dem Wissen zur Untersuchung, an einer experimentellen Überprüfung von potentiellen Einflussfaktoren, wie Informationsquellen, auf die Leistung von Entwicklern teilzunehmen. Ebenso war diesen bekannt, dass Automaten für die Prozessüberwachung und -führung eines dynamischen Systems gemeinsam in einem Team zu konzipieren sind.

Ein Untersuchungsdurchlauf startete mit der Begrüßung der drei VPn sowie einer kurzen Danksagung für die Teilnahmebereitschaft. Hiernach wurde den VPn kurz mündlich der Untersuchungszweck und -ablauf erläutert. Den VPn wurde mitgeteilt, dass jederzeit Verständnisfragen an den VI gerichtet werden können. Danach wurden den VPn eine Teamnummer von 20<sup>108</sup> bis 49 und eine Identifikationsnummer von eins bis drei zur Befestigung am Oberkörper für die Videoaufnahmen in der Reihenfolge des Erscheinens zugewiesen. Ferner wurden diese gebeten, für die gesamte Untersuchung das Handy auszuschalten. Daraufhin füllten die VPn eine Einverständniserklärung aus, in welcher die Anonymität der erhobenen Daten versichert wurde. Hiernach wurde die aktuelle Stimmung der VPn schriftlich erhoben. Dem folgte die Vorstellung der Rahmensituation für die Konzeption von Automaten für SAM (s. Anhang G.1.1). Die VPn sollten sich in die Situation hineinversetzen, ein Entwicklerteam von Automaten zu sein, welches an einem Wettbewerb zur Konzeption von Automaten zur Prozessüberwachung und -führung von dynamischen,

---

<sup>108</sup> Die Teamnummern starteten bei 20 anstelle von eins damit die ersten Teams nicht die Annahme entwickelten, noch Verbesserungsvorschläge während des Untersuchungsablaufs tätigen zu können und so diesen zu stören.



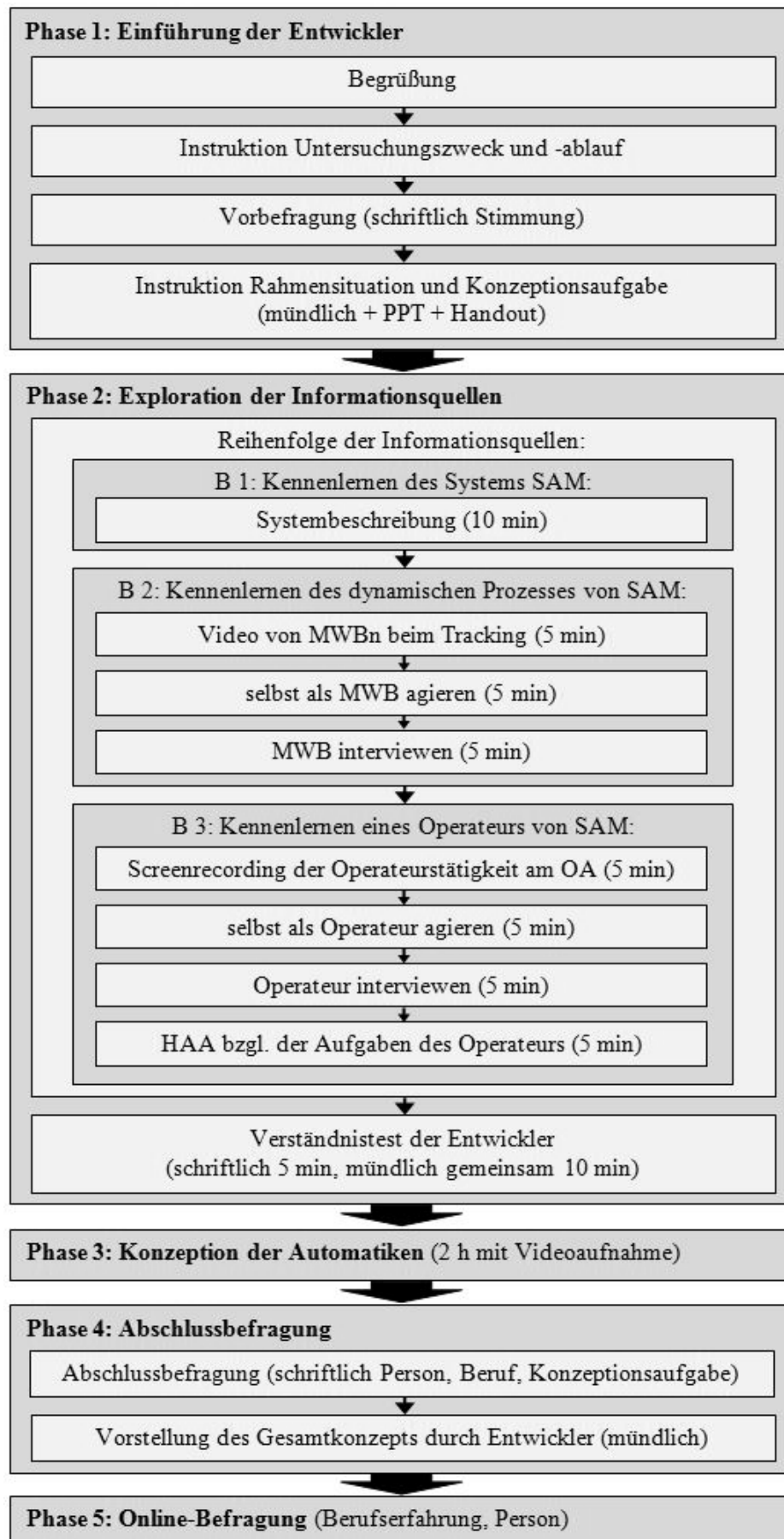


Abbildung 24: Chronologischer Ablauf der Untersuchung Leistungsdeterminante „Informationsverfügbarkeit“

technischen Systemen mit einem sehr erstrebenswerten Preisgeld teilnimmt. Um zudem die Motivation zu steigern, sollte das beste Entwicklerteam je Person einen 50 Euro Amazon-Gutschein erhalten. Danach bekamen die VPn eine kurze Beschreibung der Konzeptionsaufgabe sowie Vorstellung des Spezifikations-bogens zur Dokumentation einer Automatik samt Beispiel (s. Anhang G.1.1). Dabei wurden die VPn darauf hingewiesen, dass zur Bewertung der Güte der Konzepte nur ihre Dokumentation in den Spezifikationsbogen herangezogen wird. Anschließend wurden die Rahmenbedingungen der Konzeption vorgestellt, wie lautes Denken erwünscht sowie freie Wahl bei der Vorgehensweise der Aufgabenerfüllung.

Alle Instruktionen wurden den VPn mündlich durch den VI und parallel dazu in einer PPT-Präsentation sowie einem permanent zur Verfügung stehenden Handout präsentiert (s. Anhang G.1.1).

### ***Phase 2: Exploration der Informationsquellen***

Diese Phase begann mit dem Starten der Videoaufnahme. Danach erhielten die Teams nacheinander in einer festen Reihenfolge die einzelnen Informationsquellen (s. Abbildung 24). Dabei konnten die Teams der Bedingung 1 nur die Systembeschreibung sichten. Die Teams der Bedingung 2 bekamen zusätzlich die Informationsquellen zum Kennenlernen des dynamischen Prozesses von SAM. Bei der Selbstauführung der Objektregelung führte jede VP mit jedem Teampartner einmal das kooperative Tracking aus. Dabei nahm jeder einmal die Rolle des genauigkeitsorientierten und einmal des geschwindigkeitsorientierten MWBs ein. Die Teams der Bedingung 3 erhielten zudem die Informationsquellen zum Kennenlernen eines Operators von SAM. Damit sich die VPn nicht auf die Funktionalitäten des präsentierten OAs bei der Konzeption von Automaten konzentrieren, wurde dieser als eine der möglichen Realisierungen vorgestellt. Jede VP schlüpfte einmal in die Rolle des Operators. Währenddessen fungierten die anderen beiden VPn als MWB. Wobei jeder wieder den Schwerpunkt einmal auf schnellem und einmal auf genauem Tracking zu legen hatte.

Nachdem die VPn alle Informationsquellen erhalten hatten, wurden diese gebeten, die für die Konzeptionsaufgabe als wichtig erscheinenden Informationen aus den Informationsquellen zuerst einzeln schriftlich und danach gemeinsam mündlich zu beschreiben. Dies sollte der Erfassung des Systemverständnisses in Abhängigkeit von der Anzahl der bereitgestellten Informationsquellen dienen.

### ***Phase 3: Konzeption der Automaten***

Hiernach wurde den Teams noch einmal die Aufgabe erläutert und zu beachtende Einschränkungen bei der Konzeptionsarbeit vorgestellt (s. Anhang G.1.1). Daraufhin begann die

aktive maximal zweistündige Konzeptionsphase, in welcher die Entwickler mindestens ein Konzept von Automaten zur Prozessüberwachung und -führung von SAM 2.0 entwickeln sollten. Dabei durften sie beliebig viele Automaten mit einer beliebig hohen Anzahl an Funktionen entwickeln. Die Teams mussten selbst entscheiden, welche Automaten mit welchen Funktionen notwendig für die Überwachung und Führung des Prozesses von SAM sind, um einen optimalen Verlauf des Systemprozesses (schnelle und genaue Objektregelung) zu erreichen. Für die Dokumentation des Konzepts erhielt jedes Team einen Stapel mit leeren Spezifikationsbogen, Kugelschreiber und 20 Buntstifte. Zudem standen alle Informationsquellen bis zum Ende der Konzeptionsphase zur Verfügung.

Der VI griff in die Konzeption nur ein, wenn die Aufgabenstellung oder eine der Einschränkungen nicht beachtet wurden. Sobald ein Team sein Konzept als fertig erachtete, wurde die Konzeptionsphase beendet.

#### ***Phase 4: Abschlussbefragung***

Im Anschluss an die Konzeptionsphase erhielt jede VP einen schriftlichen Fragebogen zur Erfassung einiger Daten zu ihrer Person, beruflichen Tätigkeit, der zuvor getätigten Konzeptionsarbeit, den Informationsquellen sowie möglicher Verständnisprobleme während der Untersuchung (s. Anhang G.5). Währenddessen sichtete der VI die Spezifikationsbogen, um sich im Anschluss an den Fragebogen das Gesamtkonzept vorstellen zu lassen und dabei Verständnisfragen stellen zu können. Hiernach wurden die Videoaufnahmen gestoppt. Es folgte der Hinweis auf den Online-Fragebogen und dessen Relevanz für die Untersuchung. Am Schluss wurde darauf hingewiesen, dass in den nächsten Wochen niemandem von dem Inhalt der Untersuchung erzählt werden sollte, damit alle Entwickler mit dem gleichen Vorwissen in die Untersuchung gehen. Bei der Verabschiedung wurde den VPn für deren Teilnahme gedankt.

#### ***Phase 5: Online-Befragung***

Ein bis fünf Tage nach der Untersuchung erhielten die Entwickler eine E-Mail mit einem Link zum Online-Fragebogen, welcher allgemeine Fragen zur Berufserfahrung, Big Five und zum Leistungsmotiv umfasste (s. Anhang G.5). Wurde der Bogen nicht innerhalb von einer Woche ausgefüllt, erhielten die Entwickler eine Erinnerung per E-Mail.

### **5.2.3.2.6 Versuchsteilnehmer**

#### **5.2.3.2.6.1 Überlegungen**

Um die Konfundierung der vier Faktoren in der Untersuchung „Ressourcenmaximierung“ von Krinner (2009) und Henkel (2007) aufzulösen, sollten diese diesmal so weit

als möglich konstant gehalten werden. Eine Ausnahme stellt der Faktor „Informationsverfügbarkeit“ dar, welcher systematisch variiert werden sollte. Dafür wurde sich an den Ausprägungen der ressourcenmaximierten Bedingung angelehnt. Deshalb und aus versuchsökonomischen Gründen bestand ein Team aus drei Personen. Diese Teamgröße wurde zudem gewählt, weil sich einige Gruppenphänomene erst ab drei Personen ergeben können (Nerdinger, Blickle, & Schaper, 2011).

Ebenso sollten erfahrene Entwickler zum Einsatz kommen. Dabei mussten die Entwicklerteams aus professionellen Entwicklern und nicht aus erfahrenen Studenten bestehen. Diese Experten sollten mehrjährige Erfahrung in der Entwicklung von Automationen oder zumindest mit komplexen, dynamischen, technischen Systemen besitzen. Denn Prozesse während einer Entwicklungstätigkeit unterscheiden sich häufig nicht nur zwischen realen Arbeitssituationen und reinen Laboruntersuchungen, sondern auch zwischen Experten und Studenten (Endsley et al., 2007). Experten gehen bei der Lösung eines Problems häufig systematisch und vorausschauend vor (Ahmed, Wallace, & Blessing, 2003; Anderson, 2001; Badke-Schaub, 2005; G. Fricke, 1993). Wobei teilweise diese strukturierte Vorgehensform durch opportunistisches Verhalten unterbrochen wird (Hacker & Sachse, 2006). Laien dagegen weisen hauptsächlich ein opportunistisches Entwicklungsverhalten auf (Badke-Schaub, 2005; Christiaans & Dorst, 1992). Ebenso wirken Team-Variablen, wie die Big Five, bei Experten und Studenten unterschiedlich. So zeigen sich bspw. bei der durchschnittlichen Gewissenhaftigkeit im Team nur bei Experten ein signifikant positiver und bei der mittleren emotionalen Stabilität von Teams nur bei Studenten ein signifikant negativer Zusammenhang mit deren Leistung (Peeters et al., 2006). Somit kann nur durch erfahrene Entwickler ein hoher Praxisbezug erreicht werden (Kain, 2008). Wodurch erst allgemeingültige Aussagen getroffen werden können. Ferner sollte ein Team interdisziplinär aufgestellt sein, indem es sich aus Informatikern, Ingenieuren, Usability-Experten und Designern zusammensetzt. Diese Disziplinen sind am häufigsten in der Entwicklung von technischen, dynamischen Systemen tätig (Steinheider, 2001). Die positive Wirkung interdisziplinärer Zusammenarbeit konnte bereits in anderen Untersuchungen gezeigt werden (Anderson, 2001; Dorst, 2003; Weck, 2006).

#### **5.2.3.2.6.2 Akquise**

Es wurde eine intensive Akquise von professionellen Entwicklern von Assistenzsystemen und Automaten über einen Zeitraum von ein und einviertel Jahr betrieben. Dafür wurden Akquisetätigkeiten direkt auf themenrelevanten Tagungen, Kongressen und Messen, wie bspw. der AAET 09 in Braunschweig, per Telefonat und E-Mail unternommen. Die Gewinnung von Entwicklerteams gestaltete sich sehr aufwändig. Ein Unternehmen musste drei

Personen gleichzeitig für mindestens einen halben Arbeitstag für die Untersuchung freistellen. Dies stellt unter dem hohen zeitlichen Entwicklungsdruck der meisten Unternehmen eine große Entbehrung dar. Eine finanzielle Entlohnung für die Teilnahme konnte aufgrund des Projektbudgets nicht vorgenommen werden. Jedoch wurde den Entwicklern ein individuelles, schriftliches Feedback zur Konzeptionstätigkeit und den erhobenen Variablen angeboten, was 53 der Entwickler in Anspruch nahmen.

#### **5.2.3.2.6.3 Beschreibung der Stichprobe**

Für jede Stufe der uV „Informationsverfügbarkeit“ konnten lediglich zehn Teams gewonnen werden. Ein Team bestand aus drei Personen aus dem wirtschaftlichen oder universitären Bereich der Forschung und Entwicklung von dynamischen, technischen Systemen. Somit nahmen 90 VPn (sieben Frauen, 83 Männer) an der Untersuchung teil. Die VPn eines Teams stammten aus demselben Unternehmen/Institut. Es beteiligten sich u. a. Unternehmen wie Airbus Deutschland GmbH, BMW AG, EADS Deutschland GmbH und Honda R&D Europe mit einem oder mehreren Drei-Personen-Teams (s. Anhang G.6.1). 24 der Teams bestehen nur aus männlichen VPn (s. Tabelle 26). In Bedingung 1 gibt es ein gemischtes Team, bestehend aus einer Frau und zwei Männern. In Bedingung 2 und 3 existieren jeweils zwei Teams mit dieser Geschlechterverteilung. In Bedingung 3 kommt zudem ein Team vor, welches sich aus zwei Frauen und einem Mann zusammensetzt. Die VPn waren zwischen 23 und 61 Jahre alt ( $M = 36.06$ ,  $SD = 8.01$ ). Im Mittel unterschied sich das Durchschnittsalter der Mitglieder eines Teams der Bedingung 2 signifikant von den anderen beiden Bedingungen (s. Tabelle 27). In allen Stichproben bilden den größten Anteil mindestens diplomierte Ingenieure (s. Tabelle 26). In Bedingung 1 und 3 stellen studierte Informatiker den zweitgrößten Anteil dar. Weitere Hochschulabschlüsse sind seltener vertreten, wobei Designer nur einmal in Bedingung 2 und 3 und ein Diplom-Psychologe nur einmal in Bedingung 3 vorliegen. Wenn auch die gewünschte Interdisziplinarität nicht erreicht werden konnte, so kommt diese Verteilung, nach der mittlerweile gewonnenen Erfahrung der Autorin, an die reale Verteilung dieser Disziplinen bei der Entwicklung von Automationen für komplexe, dynamische Systeme heran. Ebenso macht den größten Anteil der beruflichen Tätigkeiten eine Ingenieursposition aus (s. Tabelle 26), wie bspw. System- oder Entwicklungsingenieur. Wissenschaftliche Mitarbeiter stellen die zweitgrößte Gruppe dar, von denen, bis auf eine Ausnahme, alles diplomierte Ingenieure oder Informatiker sind. Mit Ausnahme von zwei Personen, zu denen keine Angaben vorliegen, waren alle VPn im Bereich der Entwicklung, Bewertung oder Wartung von komplexen und/ oder dynamischen Systemen berufstätig. Die Stichproben der drei Bedingungen unterscheiden sich nicht signifikant und praktisch bedeutsam hinsichtlich der

Tabelle 26: Geschlechterverteilung in den Teams, Ausbildung/Hochschulabschluss und berufliche Position der VPn der drei Bedingungen der uV „Informationsverfügbarkeit“ und für die Gesamtstichprobe in Form von Häufigkeit

	H			
	B 1	B 2	B 3	G
<b>Geschlechterverteilung im Team</b>				
3 x männlich	9	8	7	24
3 x weiblich	0	0	0	0
1 x weiblich, 2 x männlich	1	2	2	5
2 x weiblich, 1 x männlich	0	0	1	1
<b>Ausbildung/Hochschulabschluss</b>				
Hochschulabschluss Industrie Design	0	1	1	2
Hochschulabschluss Informatiker	6	1	9	16
Hochschulabschluss Ingenieur	18	23	17	58
Hochschulabschluss Mathematik/Informatik	1	0	0	1
Hochschulabschluss Physiker	4	0	1	5
Hochschulabschluss Psychologe	0	0	1	1
Habilitation Biochemie	0	1	0	1
Nachrichtentechniker	1	0	0	1
Fachhochschule	0	2	0	2
keine Angabe	0	2	1	3
<b>berufliche Position</b>				
Angestellter/Berater für IT	1	2	0	3
Assistent Manager	1	0	0	1
Entwickler	1	0	0	1
Führungsposition im Entwicklungsbereich	2	9	3	14
HMI-Experte	1	1	2	4
Ingenieur	17	7	10	34
Lehre Kommunikationselektronik	0	1	0	1
Software-/Systementwickler	3	4	4	11
wissenschaftlicher Mitarbeiter	3	6	10	19
keine Angabe	1	0	1	2

Anmerkungen. H = Häufigkeit; B 1 = Bedingung 1: nur Systembeschreibung; B 2 = Bedingung 2: Bedingung 1 + Perspektive auf System; B 3 = Bedingung 3: Bedingung 1 + 2 + Perspektive Operateur; G = Gesamt

Teammittelwerte der nachfolgenden Selbstbeurteilungen (s. Tabelle 27). Die Teams hielten sich für überwiegend kompetent bei der Entwicklungstätigkeit. Der prozentuale Anteil im beruflichen Alltag der Entwickler, den die Problemidentifizierung bis zur Konzeption einnimmt, lag bei über 60 %. Die Teammitglieder waren untereinander mittelmäßig bis eher vertraut<sup>109</sup>. Deren Stimmung<sup>110</sup> während der Untersuchung war eher gut.

<sup>109</sup> Für jedes Team wurde der Vertrautheitsindex als Mittelwert der Teammitgliederwerte aus den nachstehenden beiden Fragen gebildet: Wie bekannt waren die anderen beiden Teammitglieder und wie oft wurde bereits zusammengearbeitet?

<sup>110</sup> Hier wurde eine zweifaktorielle Varianzanalyse mit Messwiederholung trotz teilweise nicht normalverteilter Daten berechnet. Neben dem Gruppierungsfaktor „Bedingung“ (3fach gestuft) ging der Messwiederholungsfaktor „Stimmung“ (2fach gestuft: Stimmung Prä und Post) mit in die Rechnung ein. Es existiert kein äquivalentes non-parametrisches Verfahren zur

Tabelle 27: Mittelwert und Standardabweichung der Teamwerte (Mittelwert der Mitgliederwerte) einiger Kontrollvariablen pro Bedingung der uV „Informationsverfügbarkeit“ sowie Ergebnisse der statistischen Analysen (Varianzanalyse, H-Test von Kruskal-Wallis, t-Test für unabhängige Stichproben und Mann-Whitney-U-Test mit Effektgrößen und Teststärken) zur Überprüfung der Gleichheit der Stichproben dieser drei Bedingungen

M der Team-M			SD der Team-M			Signifikanztest <sup>111</sup> , Effektgröße, Teststärke
B 1	B 2	B 3	B 1	B 2	B 3	
Alter						
35.17	39.67	33.33	5.63	5.55	5.03	H(2) = 5.787, <b>p</b> = .055 <sup>112</sup> , $\eta^2_p$ <sup>113</sup> = .212, 1 - $\beta$ = .884
						B 1 vs. 2: U = 25.50, <b>p</b> <sub>zweiseitig</sub> <sup>114</sup> = .065, r = -.418, 1 - $\beta$ = .437
						B 1 vs. 3: U = 40.50, <b>p</b> <sub>zweiseitig</sub> = .492, r = -.161, 1 - $\beta$ = .138
						B 2 vs. 3: U = 21.50, <b>p</b> <sub>zweiseitig</sub> = .030, r = -.483, 1 - $\beta$ = .743
Kompetenz bei der Entwicklungstätigkeit (Skala: 1 = gar nicht bis 5 = völlig)						
3.97	3.81	3.90	0.46	0.24	0.47	H(2) = 0.091, p = .958, $\eta^2_p$ = .006, 1 - $\beta$ = .227
prozentualer Anteil Problemidentifizierung bis Konzeption im Berufsalltag <sup>115</sup>						
66.13	62.37	67.71	11.05	11.08	8.63	F(2, 22) = 0.602, p = .557, $\eta^2_p$ = .052, 1 - $\beta$ = .391
Vertrautheitsindex (Skala: 1 = gar nicht bis 4.5 völlig)						
2.93	2.78	3.04	0.66	0.77	0.88	H(2) = 0.910, p = .635, $\eta^2_p$ = .021, 1 - $\beta$ = .295
Stimmung (Kunin-Skala: 1 = negativstes Emoticon bis 7 = positivstes Emoticon)						
Stimmung Prä						Haupteffekt: F(2, 27) = 0.898, p = .419, $\eta^2_p$ = .062, 1 - $\beta$ = .548
5.30	5.33	5.43	0.60	0.44	0.16	
Stimmung Post						Interaktionseffekt: F(2, 27) = 0.283, p = .756, $\eta^2_p$ = .021, 1 - $\beta$ = .534
5.13	4.90	4.81	0.57	0.79	0.66	
Erfahrungsindex (Skala: 1 = keine bis 5 = sehr viel)						
2.53	2.21	2.70	0.71	0.39	0.44	H(2) = 5.578, <b>p</b> = .061, $\eta^2_p$ = .140, 1 - $\beta$ = .732
						B 1 vs. 2: t(18) = 1.243, <b>p</b> <sub>zweiseitig</sub> = .234, r = .315, 1 - $\beta$ = .261
						B 1 vs. 3: t(18) = 0.657, <b>p</b> <sub>zweiseitig</sub> = .521, r = .168, 1 - $\beta$ = .119
						B 2 vs. 3: t(18) = -2.655, <b>p</b> <sub>zweiseitig</sub> = .016, r = .530, 1 - $\beta$ = .751
Wichtigkeit der Teamleistung (Skala: -50 = unwichtig bis 50 = sehr wichtig)						
31.33	29.83	21.87	11.37	9.70	11.27	H(2) = 3.245, <b>p</b> = .197, $\eta^2_p$ = .141, 1 - $\beta$ = .734
						B 1 vs. 2: t(18) = 0.317, <b>p</b> <sub>zweiseitig</sub> = .755, r = .075, 1 - $\beta$ = .080
						B 1 vs. 3: U = 25.00, <b>p</b> <sub>zweiseitig</sub> = .060, r = -.423, 1 - $\beta$ = .462
						B 2 vs. 3: U = 35.00, <b>p</b> <sub>zweiseitig</sub> = .271, r = -.254, 1 - $\beta$ = .968

mehrfaktoriellen Varianzanalyse. Ein nicht signifikantes Ergebnis unter Verletzung der Voraussetzungen würde sich auch bei gegebener Normalverteilung und Varianzhomogenität zeigen (Zöfel, 2003).

<sup>111</sup> Es konnten keine signifikanten Zusammenhänge zwischen den Variablen gefunden werden, bei welchen die Gleichheit zwischen den drei Bedingungen statistisch überprüft werden sollte. Daher wurde anstatt einer multivariaten Varianzanalyse jeweils ein univariates Verfahren für die, in diesem Unterabschnitt vorgestellten, Variablen gewählt. Weitere Verfahrensauswahlbegründungen s. Abschnitt 5.2.3.3.2.

<sup>112</sup> Zwischen den Stichproben der drei Bedingungen sollten keine Unterschiede bestehen, welche zu einer unerwünschten Beeinflussung der Untersuchungsergebnisse geführt haben können. Daher wurde für die statistische Überprüfung der Gleichheit der Stichproben ein strenges Vorgehen durch Heraufheben der Signifikanzschranke auf  $p = .20$  gewählt.

<sup>113</sup> Konventionen für die Effektgröße  $\eta_p^2$  nach Cohen (1988, zitiert nach Sedlmeier & Renkewitz, 2008):  $\eta_p^2 = .01$  ist ein kleiner,  $\eta_p^2 = .06$  ist ein mittlerer und  $\eta_p^2 = .14$  ist ein großer Effekt.

<sup>114</sup> Aufgrund der Alpha-Fehler-Kumulierung durch die drei Einzelvergleiche wurde die Schranke der Signifikanz bei  $p = .067$  anstatt bei .20 gesetzt.

<sup>115</sup> In diese Auswertung gingen zwei Teams der Bedingung 1 und 3 und ein Team der Bedingung 2 nicht mit ein, da einzelne Mitglieder Werte von über 100 % erreichten.

*Anmerkungen.* M = Mittelwert; SD = Standardabweichung; B 1 = Bedingung 1: nur Systembeschreibung; B 2 = Bedingung 2: Bedingung 1 + Perspektive auf System; B 3 = Bedingung 3: Bedingung 1 + 2 + Perspektive Operateur; signifikante empirische Werte sind kursiv und fett markiert; Prä = vor der Konzeptionstätigkeit; Post = nach der Konzeptionstätigkeit

Allerdings liegt zwischen Bedingung 2 und 3 ein signifikanter und praktisch sehr bedeutsamer Unterschied bei den Teammittelwerten bzgl. der selbst angegebenen Erfahrung<sup>116</sup> mit der Entwicklung von Automationen vor. Im Mittel liegt die Erfahrung der Entwicklerteams der Bedingung 2 in Richtung eher wenig und der Bedingung 3 in Richtung eher mittelmäßig.

Ebenso besteht zwischen Bedingung 1 und 3 bei den Teammittelwerten der Selbsteinschätzung, wie wichtig es war, ein gutes Teamergebnis zu erzielen, ein signifikanter Unterschied mit einer hohen Effektgröße. Dabei liegt die Beurteilung in Bedingung 1 und 2 an der Stufengrenze zwischen „überwiegend bis völlig wichtig“, wohingegen bei Bedingung 3 diese eindeutig als „überwiegend wichtig“ beurteilt wurde.

Aus diesen Stichprobendaten lässt sich schließen, dass es sich bei den gewonnenen VPn überwiegend um die anvisierte Zielgruppe von erfahrenen Entwicklern von komplexen, dynamischen, technischen Systemen handelt. Zudem hatte über die Hälfte der VPn bereits eine Automation entwickelt ( $H_{B1} = 18$ ,  $H_{B2} = 21$ ,  $H_{B3} = 14$ ). Ein Drittel der VPn machte zu dieser Frage keine Angaben (s. Anhang G.6.2).

Die bestehenden signifikanten und praktisch bedeutsamen Unterschiede zwischen den Stichproben werden bei der weiteren Datenanalyse berücksichtigt. Weitere Details zu den Stichproben sind dem Anhang G.6.2 zu entnehmen.

### 5.2.3.3 Ergebnisse

Dieses Kapitel dient dazu, die Aufbereitung der erhobenen Daten, deren Analyse sowie die so gewonnenen Ergebnisse in der Reihenfolge der Hypothesen vorzustellen. Zunächst wird die Aufbereitung der objektiven, dann der subjektiven Daten erläutert. Dem folgt die Präsentation der Datenanalyse inklusive der Ergebnisse.

#### 5.2.3.3.1 Aufbereitung der Daten

Für die deskriptiven und inferenzstatistischen Analysen wurden die Daten der aufgaben- und sozio-emotionalbezogenen Prozessvariablen aus den Videoaufnahmen und die Teamleistung bei der Konzeption aus den Spezifikationsbogen gewonnen. Zugleich wurden die Frage-

---

<sup>116</sup> Für jedes Team wurde der Erfahrungsindex als Mittelwert der Teammitgliederwerte zu den folgenden fünf Fragen berechnet: 1. Beschäftigung im Privatleben mit der Thematik „Mensch-Maschine-Funktionsteilung“ bzw. 2. mit „Automatisierung“, 3. Erfahrung mit der Entwicklung von Assistenzsystemen und Automaten zur Prozessüberwachung und -führung von dynamischen Systemen respektive 4. komplexen Systemen und 5. komplexen, dynamischen Systemen.



bogendaten für die Analyse vorbereitet. Aufgrund der erfassten Entwicklerleistung auf Teamebene, wurden die individuellen Werte aller weiteren erhobenen Variablen jeweils zu einem Teammittelwert aggregiert. Alle aufbereiteten Daten wurden in eine gemeinsame XLS-Datei überführt. Wobei jede Zeile einem Team und jede Spalte einer Variable entsprach. Nachfolgend werden diese Aufbereitungen kurz dargestellt.

#### 5.2.3.3.1.1 Objektive Konzeptionsdaten

Für die Datenerfassung des Großteils der Prozessvariablen wurden die Videoaufnahmen von der Konzeptionsphase mittels des IKDs (Schermuly, 2007, 2010; Schermuly & Scholl, 2010; Schermuly et al., 2010) durch Dodt (2012) im Rahmen einer Diplomarbeit, die von der Autorin betreut wurde, ausgewertet. Bei drei Teams der Bedingung 3 existiert aufgrund von technischen Problemen bei der Videospeicherung keine Tonspur, weshalb deren Entwicklungsprozess nicht mit erhoben wurde.

Zur Erfassung der *aV* „**Quantität der Automatikkonzepte**“ wurden die, von den Entwicklern konzipierte, Anzahl an Automaten und an Funktionen nach den Spezifikationsbogen<sup>117</sup> manuell bestimmt und digitalisiert. Daraus wurde die durchschnittliche Anzahl an Funktionen je Automat berechnet. Ebenso wurde die Anzahl an konzipierten Funktionen nach der Bewertungsfunktionsmatrix<sup>118</sup> manuell ermittelt und digitalisiert.

Die *aV* „**Qualität der Automatikkonzepte**“ wurde anhand der Spezifikationsbogen mit dem „Bewertungsinstrument der Güte von Automatikkonzepten“ (s. Abschnitt 5.2.2) bestimmt. Für eine Gleichgewichtung der beiden Instrumentkomponenten wurde pro Team die erreichte Punktzahl in der Bewertungsfunktionsmatrix und im -fragebogen zu einem Prozentwert umgewandelt, durch Relativierung an der jeweils höchsten erreichten Punktzahl der 30 Automatikkonzepte. Aus den Prozentwerten der beiden Instrumentkomponenten wurde je Team ein Mittelwert berechnet. Dadurch ergab sich für die 30 Automatikkonzepte eine quasi-intervallskalierte Bewertungsordnung der Güte der Automatikkonzepte.

---

<sup>117</sup> Der, von der Autorin erstellte, Spezifikationsbogen diente den Entwickler-VPn in dieser Untersuchung zur Dokumentation der konzipierten Automaten und deren Funktionen (s. Abschnitt 5.2.3.2.1.2).

<sup>118</sup> Die Bewertungsfunktionsmatrix ist eine Komponente des, von der Autorin entwickelten, „Bewertungsinstruments der Güte von Automatikkonzepten“ (s. Abschnitt 5.2.2).

### 5.2.3.3.1.2 Subjektive Fragebogendaten

Die zwei Datensätze der Vor- und Abschlussbefragung per Papier wurden manuell digitalisiert.

Für die Online-Nachbefragung lag die Rücklaufquote bei 100 %, wobei 16 der Teilnehmer an den Fragebogen per E-Mail erinnert werden mussten. Aus dem Online-Umfrage Tool Unipark konnten die Daten in eine XLS-Datei exportiert werden.

### 5.2.3.3.2 Analyse der Daten

Die Auswertung der aufbereiteten Daten erfolgte mittels MS Office Excel 2007, SPSS 19, 23 und GPower 3.1. Die deskriptiven Kennwerte (M, SD) zur Aufklärung der Forschungsfragestellung 5 sind dem Anhang G.6.4 und G.6.6 zu entnehmen. Für die inferenzstatistischen Absicherungen der Ergebnisse wurde das Alpha-Fehler-Niveau a priori auf 5 % festgelegt. Bei Vorhandensein von Ausreißern<sup>119</sup> wurden die deskriptiven und inferenzstatistischen Analysen mit und ohne Ausreißer durchgeführt. Dieses Vorgehen wurde gewählt, um Messfehler in den Daten aufgrund der Erfassungsmethoden auszuschließen.<sup>120</sup> Aufgrund der kleinen Teamstichprobe je Bedingung der Untersuchung kann nicht eindeutig geklärt werden, ob es sich tatsächlich um Ausreißer handelt. Zudem sollten diese kleinen Stichproben nicht weiter dezimiert werden. Daher werden nachfolgend sämtliche Daten inklusive der Ausreißer berichtet. Ergaben die statistischen Überprüfungen unter Ausschluss der Ausreißer ein bedeutsam anderes Ergebnis, wird dies ebenfalls berichtet. Im Anhang G.6.6 sind die Ergebnisse der Analysen exklusive Ausreißer nachzulesen.

Für die Überprüfung der einzelnen Hypothesen wurden verschiedene statistische Verfahren verwendet. Für die Unterschiedsüberprüfung zwischen den drei Bedingungen pro relevanter Variable wurde eine univariate einfaktorielle Varianzanalyse gerechnet, wenn deren Voraussetzungen, Normalverteilung pro Stichprobe und homogene Fehlervarianzen, vorlagen.<sup>121</sup> Waren deren Voraussetzungen nicht erfüllt, wurde der H-Test von Kruskal-

---

<sup>119</sup> Ausreißeranalysen erfolgten bei normalverteilten Variablen mittels einer z-Wert-Analyse und bei nicht normalverteilten Variablen über den Boxplot.

<sup>120</sup> Um im Online-Umfrage Tool Unipark von einer zur nächsten Seite blättern zu können, mussten alle Fragen der offenen Seite beantwortet werden. Bei den Radio-Buttons und Checkboxes existierte keine Vorauswahl, die Schieberegler befanden sich auf Nullstellung. Damit sind ungewollte Antworten weitestgehend auszuschließen. Für die sichere Anwendung des IKDs erhielten Dodt (2012) und die Autorin ein Training durch den Instrumententwickler Schermuly. Das „Bewertungsinstrument der Güte von Automatikkonzepten“ wurde durch die Autorin selbst entwickelt. Es wurden mehrere Bewertungsschleifen durchgeführt, so dass von keiner Fehlanwendung ausgegangen wird.

<sup>121</sup> Nach Bortz (1999) ist die Varianzanalyse nur bei gleichgroßen Stichproben mit mehr als zehn Stichprobenwerten relativ robust gegenüber Verletzungen von deren Voraussetzungen.

Wallis<sup>122</sup>, also das non-parametrische Pendent gerechnet.<sup>123</sup> Ein monotoner Trend wurde aufgrund nicht gegebener äquidistanter Abstände<sup>124</sup> zwischen den drei Ausprägungen der uV mittels des Jonckheere-Terpstra-Tests<sup>125</sup> statistisch überprüft. Bei einer Overall-Signifikanz wurden Einzelvergleiche zwischen den Bedingungen mittels des t-Tests für unabhängige Stichproben durchgeführt. Voraussetzung dafür ist eine Normalverteilung der Messwerte beider Stichproben bzgl. der relevanten Variable. Ist diese Voraussetzung nicht gegeben, wurde das non-parametrische Pendant eingesetzt, der Mann-Whitney-U-Test. Die Überprüfung eines Zusammenhangs zwischen zwei Variablen erfolgte bei gegebener Normalverteilung der Variablendaten durch den Korrelationskoeffizienten nach Pearson. Lag bei mindestens einer der Variablen keine Normalverteilung vor, wurde der Rangkorrelationskoeffizient tau von Kendall berechnet.<sup>126</sup> Häufigkeitsunterschiede zwischen den Bedingungen wurden mittels des exakten  $\chi^2$ -Tests nach Fisher statistisch überprüft, da in mehr als 20 % der Fälle die erwarteten Häufigkeiten der Zellen kleiner 5 waren.<sup>127</sup> Der unerwünschte Einfluss von Variablen auf eine Variable wurde aus dieser durch Bildung eines nicht-standardisierten Residuums mittels linearer Regressionsrechnung herauspartialisiert.

#### 5.2.3.3.3 Ergebnisse der Analysen

Nachfolgend werden die Ergebnisse zur Aufklärung der Forschungsfragestellung 5 im Detail dargestellt. Zunächst soll jedoch kurz auf die dafür relevante Analyse einiger der, in dieser Untersuchung erhobenen, Kontrollvariablen eingegangen werden. Diese Analyse diente

---

<sup>122</sup> Da keine Methoden existieren, um die Effektgröße und Teststärke bei der Anwendung des H-Tests nach Kruskal-Wallis zu berechnen, wurden zur Bestimmung dieser Werte die Methoden für das Verfahren der einfaktoriellen Varianzanalyse angewendet. Durch dieses Vorgehen, welches von B. Rasch, Hofmann, Friese und Naumann (2010) empfohlen wird, wird eine approximative Schätzung der Teststärke gewonnen, welche jedoch von der tatsächlichen Teststärke abweichen kann. Dieses Vorgehen wurde bei allen durchgeführten H-Tests nach Kruskal-Wallis in dieser Dissertation angewendet und wird daher nachstehend nicht erneut erwähnt.

<sup>123</sup> Für die Überprüfung der Normalverteilung wurden sowohl der Kolmogorov-Smirnov-Test mit Lillifors-Schranken als auch der Shapiro-Wilk-Test verwendet. Der Shapiro-Wilk-Test ist für Stichproben unter 50 Personen ausgelegt und weist für solche Stichprobengrößen eine hohe Teststärke auf (Seier, 2002). Die Normalverteilung einer Variable wurde nur angenommen, wenn dies durch beide Verfahren mit einem nicht signifikanten Ergebnis ( $p > .20$ ) statistisch bestätigt werden konnte. Die Varianzhomogenität wurde mittels des Levene-Tests überprüft. Diese wurde nur angenommen, wenn diese durch ein nicht signifikantes Ergebnis ( $p > .20$ ) statistisch belegt wurde.

<sup>124</sup> Dies stellt eine der Voraussetzungen für eine Trendanalyse mittels einer Varianzanalyse dar (Bortz, 1999, 2005).

<sup>125</sup> Da keine Methode existiert die Teststärke bei der Anwendung des Jonckheere-Terpstra-Tests zu berechnen, wurde zur Bestimmung dieses Wertes die Methode für das Verfahren der einfaktoriellen Varianzanalyse verwendet. Durch dieses Vorgehen, welches von B. Rasch et al. (2010) empfohlen wird, wird eine approximative Schätzung der Teststärke gewonnen, welche jedoch von der tatsächlichen Teststärke abweichen kann.

<sup>126</sup> Dieses Verfahren ist geeigneter als die Rangkorrelation nach Spearman, da nicht von äquidistanten Merkmalsabständen zwischen den aufeinanderfolgenden Rangplätzen ausgegangen wird. Zudem ist dieses Verfahren nicht ausreißersensibel. (Bortz & Lienert, 1998)

<sup>127</sup> Um einen  $\chi^2$ -Test verwenden zu können, müssen nach Cochran (1995, zitiert nach Bortz & Lienert, 1998) 80 % der Zellen eine erwartete Häufigkeit  $> 5$  aufweisen. Ist diese Bedingung nicht erfüllt, stellt der exakte  $\chi^2$ -Test nach Fisher ein alternatives Analyseverfahren dar (Field, 2005).

zudem der Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme“ mit der Output-Variable „Teamleistung“ (s. Abschnitt 5.2.1.1), um ein Modell speziell für Entwickler von Automationen zu gewinnen.

#### **5.2.3.3.1 Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“ mit der Output-Variable „Teamleistung“**

Vorweg werden die Analyseergebnisse von Dodt (2012) wiedergegeben, die im Rahmen ihrer Diplomarbeit ermittelt wurden. Angeleitet durch die Autorin, überprüfte Dodt (2012) mittels Korrelationsanalysen die Einflüsse einiger Input- sowie Prozessvariablen auf das Konzeptionsergebnis. Für die Input-Variablen wurde der direkte, aber auch der indirekte, über den Konzeptionsprozess vermittelte, Einfluss ermittelt.

Dafür wurde aus den Prozessvariablen sowie aus der Output-Variable „Güte der Automatikkonzepte“ der Einfluss der uV „Informationsverfügbarkeit“, durch Erstellung nicht-standardisierter Residuen mittels multipler linearer Regressionsanalysen, herauspartialisiert, um durch die uV verursachte Scheinzusammenhänge zu vermeiden. Dabei ging die dreistufige uV mittels zwei dichotomer Dummy-Variablen mit Bedingung 1 als Referenzkategorie in die Rechnung ein. Dem Anhang G.4 kann entnommen werden, welche Variablen analysiert wurden.

Dodt (2012) konnte für einige Variablen (tendenziell) signifikante Zusammenhänge ermitteln. Tabelle 28 zeigt die signifikanten Ergebnisse sowie die tendenziell signifikanten Ergebnisse für die Variable „Vertrautheitsindex“. Weitere, jedoch nur zweiseitig geprüfte, tendenziell signifikante Zusammenhänge, welche mittlere Effektgrößen aufweisen, finden sich im Anhang G.6.5.<sup>128</sup> Diese Ergebnisse konnten von der Autorin durch eine eigene Analyse der, in der Entwickleruntersuchung erhobenen, Rohdaten bestätigt werden. Einzig der von Dodt (2012) beschriebene tendenziell signifikante positive Zusammenhang zwischen dem Residuum Anzahl inhaltlicher Vorschläge mit dem Residuum Güte der Automatikkonzepte muss korrigiert werden. Die Autorin fand keinen Zusammenhang zwischen diesen beiden Variablen ( $r = -.009$ ,  $p = .483$ ,  $1 - \beta = .055$ ).<sup>129</sup>

---

<sup>128</sup> Diese Zusammenhänge spielen eine untergeordnete Rolle im Rahmen der Ergebnispräsentation dieser Arbeit, daher sind diese dem Anhang G.6.5 zu entnehmen.

<sup>129</sup> Es scheint sich um einen Ablesefehler aus der Korrelationsmatrix zu handeln. Dieser, von Dodt (2012) dargestellte, tendenzielle Zusammenhang wurde von der Autorin zwischen dem Residuum Anzahl inhaltlicher Vorschläge und dem Residuum Inhalt gefunden ( $r = .298$ ,  $p = .065$ ,  $1 - \beta = .473$ ).

Tabelle 28: Produkt-Moment-Korrelationen nach Pearson zwischen Input-, Prozessvariablen und der Output-Variable zur Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“ mit der Output-Variable „Teamleistung“

Zusammenhang zwischen Teamzusammensetzung und Leistung der Entwickler					
Input-Variable	Output-Variable	r	Signifikanz	Teststärke (1 - $\beta$ )	N
Gewissenhaftigkeit	Residuum Güte der Automatikkonzepte	.467	<b>.005</b> (einseitig)	.885	30
Leistungsmotiv HE	Residuum Güte der Automatikkonzepte	.492	<b>.003</b> (einseitig)	.913	30
Vertrautheitsindex	Residuum Güte der Automatikkonzepte	.291	.059 (einseitig)	.490	30
Zusammenhang zwischen Teamzusammensetzung und Entwicklungsprozess					
Input-Variable	Prozessvariable	r	Signifikanz	Teststärke (1 - $\beta$ )	N
Leistungsmotiv FM	Residuum Elaborationtiefe	-.340	<b>.042</b> (einseitig)	.572	27 <sup>130</sup>
Leistungsmotiv FM	Residuum Anzahl aufgabenbezogener inhaltlicher Aussagen	-.411	<b>.033</b> (zweiseitig)	.612	27
Vertrautheitsindex	Residuum Positivität der sozio-emotionalen Aussagen	.290	.072 (einseitig)	.455	27
Zusammenhang zwischen Entwicklungsprozess und Leistung der Entwickler					
Prozessvariable	Output-Variable	r	Signifikanz	Teststärke (1 - $\beta$ )	N
Residuum Witzanzahl	Residuum Güte der Automatikkonzepte	.487	<b>.005</b> (einseitig)	.884	27
Residuum Positivität der sozio-emotionalen Aussagen	Residuum Güte der Automatikkonzepte	.380	<b>.025</b> (einseitig)	.667	27

Anmerkungen. r = Produkt-Moment-Korrelation nach Pearson; N = Anzahl der Teams; Residuen sind Variablen, aus deren Varianz der durch die uV „Informationsverfügbarkeit“ aufklärbare Anteil herauspartialisiert wurde; signifikante empirische Werte sind kursiv und fett markiert; die Ergebnisse sind der Arbeit von Dodt (2012) entnommen; von der Autorin wurden die Ergebnisse auf drei Dezimalstellen erweitert und zusätzlich die Teststärken bestimmt

Als nächstes sollen die, mittels Korrelationsanalysen ermittelten, Ergebnisse der Einflussüberprüfung einiger weiterer potentieller Leistungsdeterminanten von Entwicklern direkt auf das Konzeptionsergebnis, die Güte der Automatikkonzepte, berichtet werden. Hierbei handelt es sich um die Kontrollvariablen „Wichtigkeit der Teamleistung“, „Kompetenz bei der Entwicklungstätigkeit“, „Erfahrungsindex“ und „prozentualer Anteil Problemidentifizierung bis Konzeption im Berufsalltag“. Erneut wurde als Konzeptionsergebnis das, vom Einfluss der uV befreite, Residuum Güte der Automatikkonzepte verwendet. Da sich die durchschnittlichen Teamwerte zwischen den drei Bedingungen ebenfalls bei der Variable „Alter“ signifikant und

<sup>130</sup> Bei drei Teams ist kein Ton bei der Videoaufnahme des Konzeptionsprozesses vorhanden. Dadurch konnten von diesen Teams keine Prozessvariablen gewonnen werden. Somit gingen nur die Werte von 27 Teams bei den Zusammenhangsrechnungen mit Prozessvariablen ein.

praktisch bedeutsam unterscheiden, wurde für diese Variable ebenso der Zusammenhang mit diesem Residuum ermittelt. Wie die gefundenen Korrelationskoeffizienten zeigen, liegen nur zwischen den beiden Variablen „Wichtigkeit der Teamleistung“ und „Erfahrungsindex“ mit dem Residuum Güte der Automatikkonzepte signifikante praktisch bedeutsame Zusammenhänge vor (s. Tabelle 29). Die Kontrollvariable „Kompetenz bei der Entwicklungstätigkeit“ weist einen tendenziell signifikanten Zusammenhang mit dem Residuum Güte der Automatikkonzepte auf, welcher jedoch noch keinem mittleren Effekt entspricht.

Tabelle 29: Rangkorrelationen tau von Kendall zwischen Input-Variablen und der Output-Variable zur Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“ mit der Output-Variable „Teamleistung“

Zusammenhang zwischen Teamzusammensetzung und Leistung der Entwickler					
Input-Variable	Output-Variable	$\tau$	Signifikanz zweiseitig	Teststärke (1 - $\beta$ )	N
Alter	Residuum Güte der Automatikkonzepte	-.065	.617	.063	30
Wichtigkeit der Teamleistung	Residuum Güte der Automatikkonzepte	.335	<b>.010</b>	.451	30
Kompetenz bei der Entwicklungstätigkeit	Residuum Güte der Automatikkonzepte	.252	.076	.272	30
Erfahrungsindex	Residuum Güte der Automatikkonzepte	.304	<b>.019</b>	.379	30
prozentualer Anteil Problemidentifizierung bis Konzeption im Berufsalltag	Residuum Güte der Automatikkonzepte	-.064	.657	.061	25 <sup>131</sup>
Anmerkungen. $\tau$ = Rangkorrelationskoeffizient tau von Kendall; N = Anzahl der Teams; Residuen sind Variablen, aus deren Varianz der durch die uV „Informationsverfügbarkeit“ aufklärbare Anteil herauspartialisiert wurde; signifikante empirische Werte sind kursiv und fett markiert					

#### 5.2.3.3.2 Ergebnisse zur Forschungsfragestellung 5

##### 5.1 Forschungshypothesen bezüglich aV „Quantität der Automatikkonzepte“

Die explorative Analyse der Konzeptionsergebnisse erbrachte, dass jedes der 30 Teams nur ein Automatikkonzept erstellt hatte. Die *Hypothesen F.5.1.1, F.5.1.2, F.5.1.3* und *F.5.1.4* postulieren Unterschiede zwischen den drei Bedingungen hinsichtlich der, durch die Entwicklerteams konzipierten, Anzahl an Automaten, Anzahl an Funktionen und durchschnittlichen Anzahl an Funktionen je Automat nach den Spezifikationsbogen sowie Anzahl an Funktionen nach der Bewertungsfunktionsmatrix. Die statistischen Überprüfungen dieser Hypothesen erfolgten mittels des H-Tests nach Kruskal-Wallis. Als Ergebnis zeigen sich keine signifikanten und praktisch bedeutsamen Unterschiede zwischen den drei Bedingungen hinsichtlich der

<sup>131</sup> In diese Auswertung gingen zwei Teams der Bedingung 1 und 3 und ein Team der Bedingung 2 nicht mit ein, da einzelne Mitglieder Werte von über 100 % erreichten.

Anzahl an Automaten ( $H[2] = 0.734$ ,  $p = .693$ ,  $\eta^2_p = .030$ ,  $1 - \beta = .117$ ), der Anzahl an Funktionen ( $H[2] = 1.884$ ,  $p = .390$ ,  $\eta^2_p = .049$ ,  $1 - \beta = .168$ ) und der durchschnittlichen Anzahl an Funktionen je Automatik nach den Spezifikationsbogen ( $H[2] = 0.631$ ,  $p = .730$ ,  $\eta^2_p = .052$ ,  $1 - \beta = .175$ ) sowie der Anzahl an Funktionen nach der Funktionsmatrix ( $H[2] = 1.427$ ,  $p = .490$ ,  $\eta^2_p = .055$ ,  $1 - \beta = .184$ ; s. Abbildung 25).

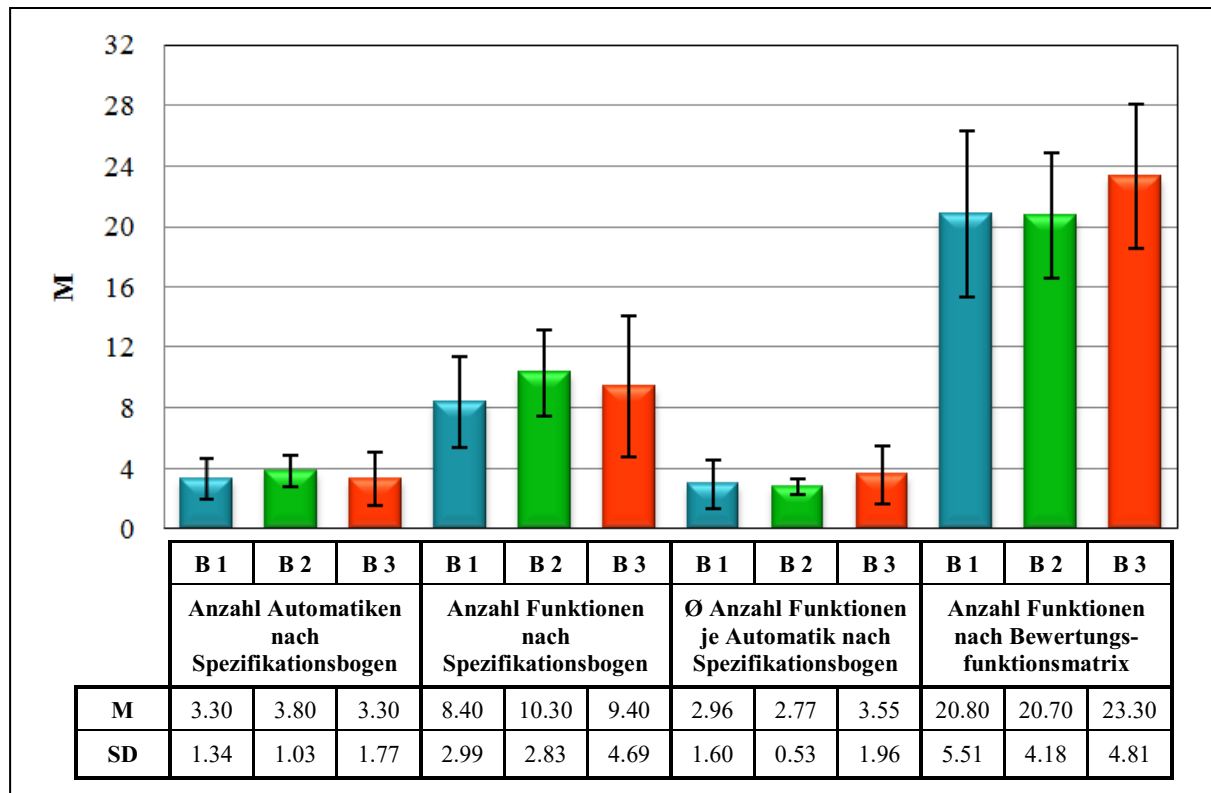


Abbildung 25: Mittelwert (M) und Standardabweichung (SD) der, durch die Entwickler konzipierten, Anzahl an Automaten, Anzahl an Funktionen und durchschnittlichen Anzahl an Funktionen je Automatik nach den Spezifikationsbogen sowie Anzahl an Funktionen nach der Bewertungsfunktionsmatrix pro Bedingung (B) der uV „Informationsverfügbarkeit“

## 5.2 Forschungshypothese bezüglich aV „Qualität der Automatikkonzepte“

Die Vielfalt der, durch die Entwickler vorgeschlagenen, Lösungen ist erheblich. Sie deckt einen großen Teil der Funktionen der Bewertungsfunktionsmatrix samt deren unterschiedlichen Auslegungsmöglichkeiten ab. Dabei weist das Automatikkonzept des Teams 37 ( $G^{132} = 92.95\%$ ) der Bedingung 3 die höchste Güte auf, gefolgt von jenen der Teams 36 ( $G = 88.47\%$ ) und 48 ( $G = 86.63\%$ ) der Bedingung 2 sowie jenen der Teams 47 ( $G = 80.66\%$ ) und 38 ( $G = 79.50\%$ ) der Bedingung 3 (s. Anhang G.6.).

<sup>132</sup> G = Güte des Automatikkonzepts

Die **Hypothese F.5.2** postuliert einen Unterschied zwischen den drei Bedingungen hinsichtlich der Qualität der Automatikkonzepte in Form eines monotonen Trends. Mit ansteigender Anzahl an Informationsquellen von Bedingung 1 über Bedingung 2 hin zu Bedingung 3 fällt die Güte der Automatikkonzepte höher aus. Der Jonckheere-Terpstra-Test ergibt für den Gruppierungsfaktor „Bedingung“ (3fach gestuft) nur einen tendenziell signifikanten monotonen Trend in die erwartete Richtung ( $J = 185.00$ ,  $z = 1.331$ ,  $p_{\text{einseitig}} = .096$ ,  $\eta^2_p = .059$ ,  $1 - \beta = .196$ ; s. Abbildung 26).

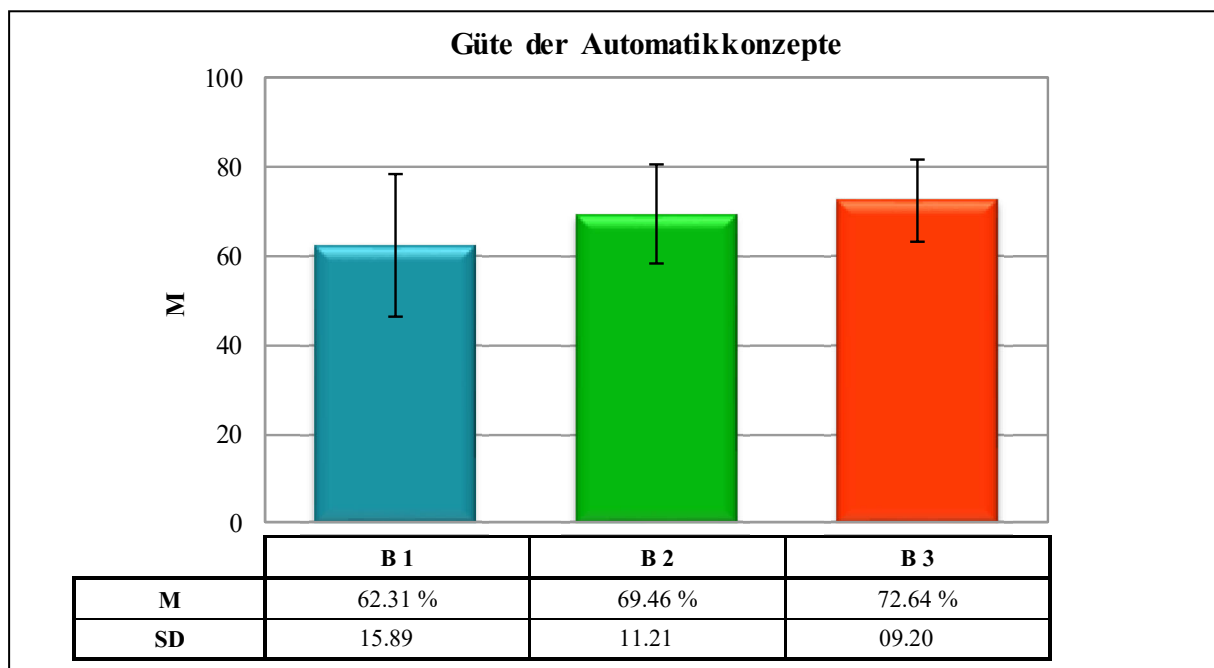


Abbildung 26: Mittelwert (M) und Standardabweichung (SD) der Güte der Automatikkonzepte pro Bedingung (B) der uV „Informationsverfügbarkeit“

### 5.3 Kontrollhypothese bezüglich des Einflusses der Kontrollvariablen auf die aV „Qualität der Automatikkonzepte“

Die **Hypothese K.5.3** postuliert eine Beeinflussung der Beziehung zwischen der uV „Informationsverfügbarkeit“ und der aV „Güte der Automatikkonzepte“ durch Kontrollvariablen. Daher wurde als nächstes überprüft, ob die zuvor ermittelte Wirkung der uV auf die aV durch einige der erhobenen Kontrollvariablen bewirkt wurde. Aufgrund der gefundenen Zusammenhänge der sechs Kontrollvariablen „Gewissenhaftigkeit“, „Leistungsmotiv HE“, „Wichtigkeit der Teamleistung“, „Erfahrungsindex“, „Residuum Witzanzahl“ sowie „Residuum Positivität der sozio-emotionalen Aussagen“ (s. Abschnitt 5.2.3.3.1) mit dem Residuum



Güte der Automatikkonzepte, wurde durch eine schrittweise<sup>133</sup> multiple lineare Regression<sup>134</sup> ein weiteres nicht-standardisiertes Residuum der aV „Güte der Automatikkonzepte“ gebildet.

Tabelle 30: Schrittweise multiple lineare Regression mit den Prädiktoren „Erfahrungsindex“, „Residuum Witzanzahl“ und „Leistungsmotiv HE“ sowie dem Kriterium „Güte der Automatikkonzepte“

	B	SE B	$\beta_i$	Sig.	Korrelation			R <sup>2</sup> /ΔR <sup>2</sup>	Sig.
					0ter Ordnung	Partiell	Teil		
Modell 1									
Konstante	36.312	9.931		.001					
Erfahrungsindex	13.045	3.970	.549	.003	.549	.549	.549	.302	.003
Modell 2									
Konstante	39.110	9.060		<.001					
Erfahrungsindex	11.962	3.620	.504	.003	.549	.559	.500		
Residuum Witzanzahl	6.481	2.544	.388	.018	.447	.461	.385	.450/ .149	.018
Modell 3									
Konstante	-4.099	18.881		.830					
Erfahrungsindex	10.231	3.338	.431	.005	.549	.539	.419		
Residuum Witzanzahl	6.212	2.299	.372	.013	.447	.491	.369		
Leistungsmotiv HE	14.485	5.705	.355	.018	.472	.468	.347	.571/ .120	.018

Anmerkungen. B = nicht-standardisierter Beta-Koeffizient; SE B = Standardfehler des nicht-standardisierten Beta-Koeffizienten;  $\beta_i$  = standardisierter Beta-Koeffizient; Sig. = Signifikanz;  $R^2$  = Bestimmtheitsmaß: Anteil der Varianz des Kriteriums, der durch den Prädiktor/die Prädiktoren vorhergesagt werden kann;  $\Delta R^2$  = Differenz des  $R^2$  zu dem des vorhergehenden Modells, verursacht durch den jeweils neuen Prädiktor; Konventionen für die Effektgröße  $R^2$  nach Cohen (1988, zitiert nach Ellis, 2010):  $R^2 = .02$  ist ein kleiner,  $R^2 = .13$  ist ein mittlerer und  $R^2 = .26$  ist ein großer Effekt; signifikante empirische Werte sind kursiv und fett markiert

Dadurch wurde diesmal der durch die sechs Kontrollvariablen signifikant aufklärbare Varianzanteil aus der aV herauspartialisiert. Mittels dieses zweiten Residuums der aV wurde

<sup>133</sup> Da es keine Annahmen aufgrund von Vorerfahrungen hinsichtlich der Vorhersagekraft der einzelnen Kontrollvariablen bzgl. der aV „Güte der Automatikkonzepte“ gab und die Korrelationskoeffizienten dicht beieinanderliegen, wurde keine hierarchische multiple lineare Regression vorgenommen. Stattdessen wurde das schrittweise Verfahren verwendet, bei dem die potentiellen Prädiktoren nach der Größe von deren mathematisch bestimmten signifikanten Vorhersagekraft in die Regressionsrechnung eingehen. (s. Field, 2005)

<sup>134</sup> Zur Absicherung der Vorhersagekraft der Kontrollvariablen als Prädiktoren für die aV „Güte der Automatikkonzepte“ als Kriterium wurde zusätzlich für jede Variable eine einfache lineare Regressionsrechnung ausgeführt. Die Vorhersagekraft eines Prädiktors hängt von den anderen Variablen einer Regressionsrechnung ab, da diese zumeist einen redundanten Informationsanteil aufweisen (Field, 2005). Die Ergebnisse der einfachen linearen Regressionen unterstützen das gefundene Ergebnis der multiplen linearen Regressionsrechnung. Die Prädiktoren „Wichtigkeit der Teamleistung“ ( $R^2 = .072$ ,  $p = .153$ ,  $1 - \beta = .438$ ) und „Residuum Positivität der sozio-emotionalen Aussagen“ ( $R^2 = .125$ ,  $p = .070$ ,  $1 - \beta = .605$ ) können die Varianz des Kriteriums nicht signifikant aufklären. Dies könnte jedoch an den zu geringen Teststärken liegen, da die jeweils aufgeklärte Varianz in Richtung eines mittleren Effekts liegt. Der Prädiktor „Gewissenhaftigkeit“ kann zwar einen Teil der Varianz des Kriteriums „Güte der Automatikkonzepte“ signifikant aufklären ( $R^2 = .195$ ,  $p = .015$ ,  $1 - \beta = .838$ ). Jedoch korreliert die Variable „Gewissenhaftigkeit“ mit den Variablen „Leistungsmotiv HE“ ( $r = .51$ ,  $p = .003$ ,  $1 - \beta = .936$ ) und „Erfahrungsindex“ ( $r = .44$ ,  $p = .012$ ,  $1 - \beta = .835$ ) signifikant. Daher trägt die Variable „Gewissenhaftigkeit“ redundante Informationen zu den Variablen „Erfahrungsindex“ und „Leistungsmotiv HE“. Diese redundante Information, welche einen Teil der Varianz der aV erklären kann, wurde somit in der multiplen linearen Regression bereits durch die letzteren beiden Variablen aufgeklärt. Dadurch ist die Information der Variable „Gewissenhaftigkeit“ zur Vorhersage der aV in dieser Regressionsrechnung überflüssig. Ebenso ist diese damit bereits aus dem gebildeten nicht-standardisierten Residuum Güte der Automatikkonzepte herauspartialisiert.

der Einfluss der uV erneut geprüft. Dieses Vorgehen wurde gewählt, weil der Jonckheere-Terpstra-Test keine direkte Aufnahme von Kontrollvariablen in die Berechnung ermöglicht.

Wie sich durch die Regressionsrechnung zeigte, kann die Variable „Erfahrungsindex“ (30 %) den höchsten Anteil der Varianz der aV „Güte der Automatikkonzepte“ aufklären, gefolgt von dem Residuum Witzanzahl (15 %) und dem Leistungsmotiv HE (12 %) (s. Tabelle 30). Die anderen drei Variablen wurden aus der Regressionsrechnung ausgeschlossen<sup>135</sup>, da sie keinen signifikanten Vorhersagebeitrag hinsichtlich der aV im Rahmen dieser Regressionsrechnung leisten können.

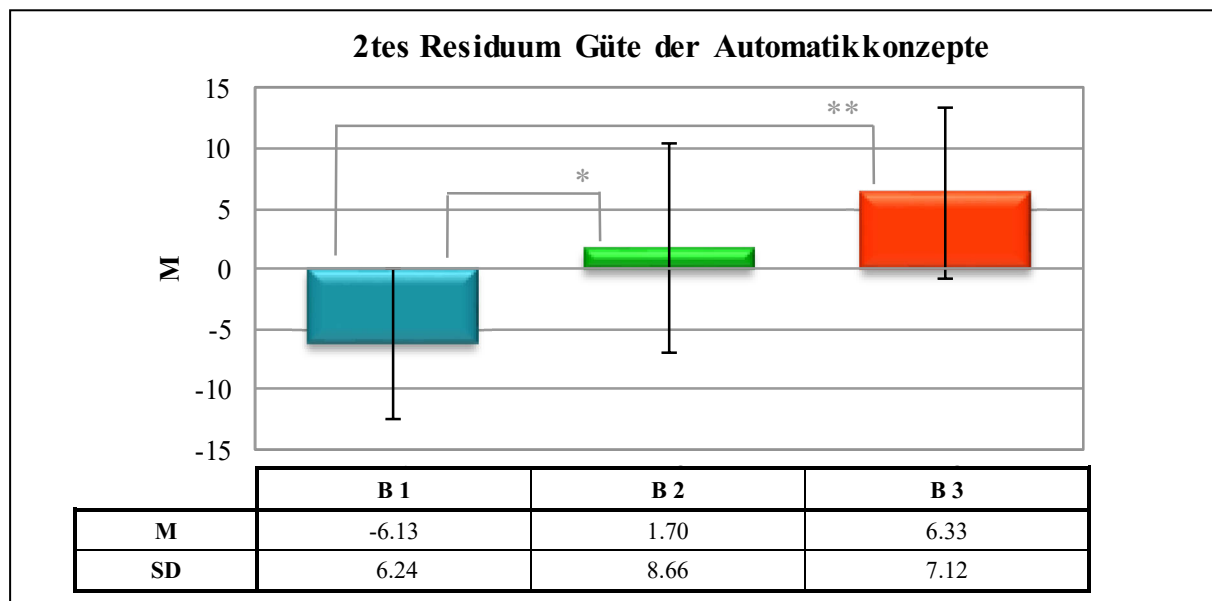


Abbildung 27: Mittelwert (M) und Standardabweichung (SD) des zweiten Residuums Güte der Automatikkonzepte (ohne gemeinsamen Varianzanteil der Variablen „Erfahrungsindex“, „Residuum Witzanzahl“ und „Leistungsmotiv HE“) pro Bedingung (B) der uV „Informationsverfügbarkeit“

Anmerkung. signifikante Unterschiede sind markiert: \* =  $p_{\text{einseitig}} < .0167$ , \*\* =  $p_{\text{einseitig}} < .003$  (Signifikanzschranken aufgrund Alpha-Fehler-Korrektur)

Für die statistische Überprüfung der Wirkung der uV „Informationsverfügbarkeit“ auf die aV „Güte der Automatikkonzepte“ unter Ausschluss der Wirkungen der Variablen

<sup>135</sup> Somit gingen drei Prädiktoren in die Regressionsrechnung ein. Nach einer verbreiteten „Daumenregel“ sollten pro Prädiktor 10 bis 15 Datenfälle in eine Regressionsrechnung einfließen (s. Field, 2005). Eine zu geringe Fallzahl kann zu einer Überschätzung des aufgeklärten Varianzanteils durch die Prädiktoren mit gleichzeitig unsicherer statistischer Überprüfung und damit zu einem nicht signifikanten Ergebnis führen (Stade, Meyer, Niestroj, & Nachtwei, 2011). Aufgrund der explorativen Analyse der Daten wird die geringfügige Unterschreitung der Forderung durch die verwendete Fallzahl von 27 Teams für die Regressionsrechnung und die damit einhergehende Gefahr in Kauf genommen.

Generell stellt diese „Daumenregel“ ein sehr liberales Vorgehen dar. Konservative Regeln fordern weitaus mehr Fälle. So stellt Green (1991, zitiert nach, Field, 2005) für die statistische Absicherung des Bestimmtheitsmaßes  $R^2$  die Formel  $N = 50 + 8 \cdot k$  ( $k$  = Anzahl Prädiktoren) auf, wonach bei drei Prädiktoren ein  $N$  von 74 benötigt wird. Selbst dieses Vorgehen stellt noch eine Vereinfachung dar. Damit der durch die Prädiktoren aufklärbare Varianzanteil des Kriteriums gut zu schätzen ist, müssten in die Bestimmung der notwendigen Fallzahl die zu erreichende Effektgröße und Teststärke mit einfließen. (s. Field, 2005)

„Erfahrungsindex“, „Residuum Witzanzahl“ und „Leistungsmotiv HE“ wurde wiederum der Jonckheere-Terpstra-Test verwendet.<sup>136</sup> Für den **Gruppierungsfaktor „Bedingung“** (3fach gestuft) zeigt sich diesmal ein höchst signifikanter monotoner Trend in die erwartete Richtung ( $J = 191.00$ ,  $z = 3.170$ ,  $p_{\text{einseitig}} = .001$ ,  $\eta^2_p = .372$ ,  $1 - \beta = .929$ ). Einzelvergleiche ergaben nur zwischen **Bedingung 1 und 2** ( $t[18] = -2.320$ ,  $p_{\text{einseitig}}^{137} = .016$ ,  $r = .480$ ,  $1 - \beta = .519$ ) sowie **1 und 3** ( $t[15] = -3.829$ ,  $p_{\text{einseitig}} = .001$ ,  $r = .703$ ,  $1 - \beta = .681$ ) nicht jedoch zwischen 2 und 3 ( $t[15] = -1.164$ ,  $p_{\text{einseitig}} = .131$ ,  $r = .288$ ,  $1 - \beta = .144$ ) einen signifikanten Unterschied (s. Abbildung 27).

#### **5.4 Explorative Analyse hinsichtlich der Unterschiede zwischen den Bedingungen der uV „Informationsverfügbarkeit“ bei den Kategorien der Bewertungsfunktionsmatrix**

Eine explorative Analyse der Automatikkonzepte sollte dazu dienen, eventuelle Unterschiede zwischen den drei Bedingungen der uV „Informationsverfügbarkeit“ hinsichtlich der Berücksichtigung der Kategorien der Funktionsmatrix aufzudecken. Es zeigt sich, dass in jeder Bedingung Funktionen für sämtliche Unterkategorien der Hauptkategorie „Funktionsklasse“ der Funktionsmatrix konzipiert wurden (s. Tabelle 31 oder Anhang G.6.6, der die Ergebnisse der statistischen Überprüfung enthält). Einzige Ausnahme stellt die Unterkategorie „Veränderung der Joystick-Empfindlichkeit“ der Funktionsklasse „Handlungsausführung“ dar. Für diese erstellten lediglich jeweils zwei Teams der Bedingungen 2 und 3 Funktionen.

Dafür berücksichtigten je acht Teams der Bedingung 1 für die beiden Unterkategorien „Informationsvermittlung“ der Funktionsklassen „Informationsakquise“ und „-analyse“ Funktionen. Nur vier Teams der Bedingung 2 und fünf Teams der Bedingung 3 sahen für die Unterkategorie „Informationsvermittlung“ der Funktionsklasse „Informationsakquise“ Funktionen vor. Dabei fällt auf, dass in Bedingung 3 kein Team eine Funktion zur Anzeige von Ist-Werten plante, bspw. der Ist-Joystick-Eingabe je MWB. Interessanterweise konzipierten nur zwei Teams eine Funktion zur Anzeige einer Pfadvorausschau, welche beide der Bedingung 3 angehörten. Für die Unterkategorie „Informationsvermittlung“ der Funktionsklasse „Informationsanalyse“ fällt der Unterschied mit sechs Teams der Bedingung 2 und sieben Teams der Bedingung 3 geringer aus. Die Bedingung 1 weist in diesen beiden Unterkategorien die höchste und die Bedingung 2 die geringste Funktions- sowie Punktzahl auf. So konzipierten bspw. sieben

---

<sup>136</sup> Da ein monotoner Trend ermittelt werden sollte, wurde sich gegen eine Regressionsrechnung mit all diesen Variablen entschieden. Die dreifach gestufte uV hätte nur als zwei dichotome Dummy-Variablen in diese Rechnung einfließen können, wodurch das Vorhandensein eines monotonen Trends nicht aufdeckbar ist. Zudem liegt eine zu geringe Fallzahl mit  $N = 27$  für eine solche Regressionsrechnung vor.

<sup>137</sup> Aufgrund der drei Einzelvergleiche wurde eine Alpha-Fehler-Korrektur nach Bonferoni durch Setzung der Signifikanzschranke auf  $p = .0167$  vorgenommen.

Teams der Bedingung 1 und nur vier Teams der Bedingung 2 Funktionen, die die MWB z. B. vor einer Hinderniskollision warnen.

*Tabelle 31: Anzahl der Teams, die Funktionen konzipierten, Mittelwert und Standardabweichung der Funktionsanzahl und der erreichten Punktzahl der Teams je Unterkategorie der Hauptkategorie „Funktionsklasse“ der Funktionsmatrix pro Bedingung der uV „Informationsverfügbarkeit“ zur explorativen Überprüfung von Unterschieden bei den Automatikkonzepten*

		Anzahl Teams, die Funktionen konzipierten			Funktionsanzahl pro Team			erreichte Punktzahl pro Team		
		B 1	B 2	B 3	B 1	B 2	B 3	B 1	B 2	B 3
Informationsakquise										
Datenerfassung	H/M	10	10	10	5.40	6.60	6.50	32.80	37.10	46.50
	SD				2.01	2.84	2.07	15.85	17.84	19.75
Informationsvermittlung	H/M	8	4	5	2.20	0.70	1.00	4.93	3.15	3.95
	SD				1.81	1.06	1.33	6.14	4.61	5.26
Informationsanalyse										
Datenberechnung	H/M	10	10	10	3.80	3.30	4.80	19.50	20.30	29.20
	SD				2.04	1.25	1.62	10.39	11.45	13.21
Informationsvermittlung	H/M	8	6	7	2.00	1.00	1.90	9.00	3.07	9.00
	SD				1.76	0.94	2.18	8.33	4.06	12.17
Entscheidungsfindung										
Entscheidung	H/M	9	9	9	2.10	3.70	2.80	9.30	18.20	12.60
	SD				1.29	2.21	1.87	5.60	12.19	11.21
Informationsvermittlung	H/M	10	10	10	2.30	2.20	2.10	11.71	10.43	7.84
	SD				0.95	1.03	1.10	5.69	6.57	4.13
Handlungsausführung										
Veränderung der Joystick-Empfindlichkeit	H/M	0	2	2	0.00	0.20	0.30	0.00	1.20	1.00
	SD				0.00	0.42	0.67	0.00	2.53	2.31
Veränderung des Objektregelungsanteils zwischen MWBn	H/M	3	4	6	0.40	0.40	0.70	1.00	2.30	2.85
	SD				0.70	0.52	0.67	2.00	2.98	3.07
Eingriff in Objektregelung	H/M	10	10	10	2.60	2.60	3.20	7.19	6.48	7.14
	SD				0.84	1.43	2.10	4.66	5.62	5.52

*Anmerkungen.* H = Anzahl an Teams, die Funktionen konzipierten; M = Mittelwert der Funktionsanzahl und der erreichten Punktzahl der Teams; SD = Standardabweichung der Funktionsanzahl und der erreichten Punktzahl der Teams; B 1 = Bedingung 1: nur Systembeschreibung; B 2 = Bedingung 2: Bedingung 1 + Perspektive auf System; B 3 = Bedingung 3: Bedingung 1 + 2 + Perspektive Operateur

Die Unterkategorie „Entscheidung“ der Funktionsklasse „Entscheidungsfindung“ wurde von je einem Team der drei Bedingungen nicht beachtet. Wobei die Teams in Bedingung 2 die größte Anzahl an Funktionen für diese Unterkategorie konzipierten und dadurch die höchste Punktzahl erreichten.

Für die Unterkategorie „Veränderung des Objektregelungsanteils zwischen den MWBn“ der Funktionsklasse „Handlungsausführung“ konzipierten nur drei Teams der Bedingung 1, aber sechs Teams der Bedingung 3 Funktionen. Entsprechend fallen die Funktions- und Punktzahl in den Bedingungen aus.

Für die restlichen vier Unterkategorien wurden von allen 30 Teams Funktionen geplant. Wobei in der Unterkategorie „Datenerfassung“ der Funktionsklasse „Informationsakquise“ und der Unterkategorie „Datenberechnung“ der Funktionsklasse „Informationsanalyse“ in der Bedingung 3 die höchste und in der Bedingung 1 die geringste Punktzahl erreicht wurde.

Für die Unterkategorie „Informationsvermittlung“ der Funktionsklasse „Entscheidungsfindung“ liegt abermals in Bedingung 1 die höchste Funktions- und Punktzahl vor. Diesmal schneidet die Bedingung 3 am schlechtesten ab. So plante in dieser Bedingung bspw. nur die Hälfte der Teams die Anzeige von Soll-Werten. Dagegen sahen je sieben Teams der anderen beiden Bedingungen Funktionen dafür vor.

Bei der Unterkategorie „Eingriff in die Objektregelung“ der Funktionsklasse „Handlungsausführung“ weist die Bedingung 3 die höchste Anzahl an Funktionen auf. Trotzdem liegt aufgrund teilweise geeigneterer Auslegungen die Bedingung 1 mit 7.19 Punkten knapp vor Bedingung 3 mit 7.14 Punkten.

Sieben Teams der Bedingung 3, aber nur fünf Teams der Bedingung 2 und vier Teams der Bedingung 1 konzipierten Funktionen für die Funktionsklasse „Handlungsausführung“ mit einer adaptiven Veränderung/Übernahme des Objektregelungsanteils. Dagegen plante nur ein einziges Team der Bedingung 2 Funktionen, bei denen die Veränderung/Übernahme des Objektregelungsanteils durch die MWB adaptierbar sind. Zudem stellten sechs Teams der Bedingung 2 Funktionen ohne konkrete Angaben zur Art und Weise der Handlungsausführung auf. In Bedingung 1 handelt es sich nur um drei und in Bedingung 3 um vier Teams, welche die Handlungsausführung nicht näher spezifizierten.

Des Weiteren konzipierten vier Teams der Bedingung 3, aber nur ein Team der Bedingung 1 und zwei Teams der Bedingung 2 Funktionen mit einer graduellen Veränderung des Objektregelungsanteils zwischen den MWBn respektive Rückgabe der Objektregelung an die MWB.

Ferner hatten in Bedingung 1 vier und in den anderen beiden Bedingungen je fünf Teams durch einzelne Funktionen eine individuelle Behandlung der MWB vorgesehen. Darunter fällt bspw. eine individuelle Vermittlung der Soll-Joystick-Eingaben. Dennoch liegt in Bedingung 3 die geringste Anzahl solcher Funktionen vor. Dadurch wurde hier auch die niedrigste Punktzahl erreicht.

Dafür haben die Teams der Bedingung 3 am häufigsten und die der Bedingung 1 am seltensten bei ihren Funktionen zwischen der Quer- und Längsregelung unterschieden, indem bspw. explizit Feedbackfunktionen zur Güte der vertikalen und horizontalen Joystick-Eingabe konzipiert wurden. Entsprechend fallen die erreichten Punktwerte durch diese Funktionen aus. Achsenspezifische Funktionen wurden von allen 30 Teams aufgestellt. Es konzipierten trotzdem je acht Teams der Bedingungen 1 und 2 sowie sechs Teams der Bedingung 3 allgemeingültige Funktionen für Unterstützungsaspekte, bei denen eine solche Achsendifferenzierung möglich wäre. Dabei weisen die Teams der Bedingung 1 die höchste und die Teams der Bedingung 3 die niedrigste Anzahl an solchen achsenundifferenzierten Funktionen auf.

Mit einer Ausnahme erwies sich keiner der zuvor beschriebenen Unterschiede zwischen den Bedingungen als signifikant (s. Anhang G.6.6). Bei der Ausnahme handelt es sich um die unterschiedliche *Anzahl an Teams* zwischen den *drei Bedingungen*, welche *Funktionen* für die Funktionsklasse „Handlungsausführung“ mit einem *fest definierten Anteil* der *Veränderung/Übernahme der Objektregelung in kritischen Situationen konzipierten* (B 1 = 8, B 2 = 3, B 3 = 8,  $\chi^2_{\text{exakt}}[2] = 6.568$ ,  $p = .041$ ,  $\omega = .467$ ,  $1 - \beta = .624$ ). Aufgrund potentieller Alpha-Fehler-Kumulierung wird dieser Unterschied nachfolgend nicht weiter betrachtet.

### 5.2.3.4 Interpretation und Diskussion der Ergebnisse

#### 5.2.3.4.1 Interpretation und Diskussion der Validierung und Erweiterung des „Modells der Leistungsdeterminanten von Entwicklern“

Wie die vorliegende Untersuchung aufzeigen konnte, stehen die Input-Variablen „Gewissenhaftigkeit“, „Leistungsmotiv HE“, „Wichtigkeit der Teamleistung“, „Erfahrungsindex“ sowie die Prozessvariablen „Witzanzahl“ und „Positivität der sozio-emotionalen Aussagen“ auf Teamebene in einem signifikanten positiven Zusammenhang mit der Output-Variable „Teamleistung“. Dabei können diese Variablen, welche als Kontrollvariablen in dieser Arbeit verwendet wurden, einen größeren Varianzanteil der  $aV$  „Güte der Automatikkonzepte“ aufklären, als die systematisch manipulierte  $uV$  „Informationsverfügbarkeit“ selbst. Die  $uV$  spielt im Vergleich zu diesen Variablen eine untergeordnete Rolle für die Güte der Automatikkonzepte. Aufgrund teilweise redundanter Informationen dieser Kontrollvariablen fällt deren Vorhersagekraft hinsichtlich der Güte der Automatikkonzepte in einer gemeinschaftlichen statistischen Analyse nur für die beiden Input-Variablen „Erfahrungsindex“ und „Leistungsmotiv HE“ sowie die Prozessvariable „Residuum Witzanzahl“ signifikant aus. Die Variable „Gewissenhaftigkeit“ steht mit den Variablen „Leistungsmotiv HE“ ( $r = .51$ ,  $p = .003$ ,  $1 - \beta =$

.936), „*Wichtigkeit der Teamleistung*“ ( $r = .32, p < .05, 1 - \beta = .564$ ) und „*Erfahrungsindex*“ ( $r = .44, p = .012, 1 - \beta = .835$ ) in einem positiven Zusammenhang. Gewissenhaftigkeit, als eines der fünf zeitlich und situativ stabilen Persönlichkeitsmerkmale (McCrae & Costa, 1990), entspricht unter anderem dem Streben nach Leistung und Kompetenz (Rammstedt & John, 2007). Daher ist anzunehmen, dass die Ausprägung der Gewissenhaftigkeit Einfluss auf die Ausprägung der anderen drei Variablen nimmt.

Die Variable „*Wichtigkeit der Teamleistung*“ korreliert ebenfalls mit den Variablen „*Erfahrungsindex*“ ( $r = .34, p = .041, 1 - \beta = .613$ ) und „*Leistungsmotiv HE*“ ( $r = .52, p = .002, 1 - \beta = .946$ ). Das Leistungsmotiv HE entspricht der zeitlich überdauernden Neigung in bestimmten Situationen (Huber et al., 2008) nach Erfolg zu streben (Lang & Fries, 2006). Durch bestimmte Bedingungen einer Situation wird aus diesem Leistungsmotiv eine konkrete Leistungsmotivation. Somit sollte sich die Ausprägung einer Person hinsichtlich der Variable „*Leistungsmotiv HE*“ in der *Wichtigkeit der Teamleistung* widerspiegeln. Ebenso könnte die Variable „*Erfahrungsindex*“ ursächlich für den gemeinsamen Varianzanteil mit der Variable „*Wichtigkeit der Teamleistung*“ sein. Denn je mehr theoretische und praktische Erfahrung die Entwickler besitzen, desto höher könnten diese ihre Chance eingeschätzt haben, die beste Teamleistung zu erbringen und somit den offerierten „Gewinn“ zu erhalten. Durch diese Einschätzung, so die Annahme, müsste sich eine höhere *Wichtigkeit der Teamleistung* ergeben. Die Variable „*Leistungsmotiv HE*“ steht in einem geringen, nicht signifikanten Zusammenhang mit der Variable „*Erfahrungsindex*“ ( $r = .21, p = .146, 1 - \beta = .310$ ). Ein Scheinzusammenhang zwischen den Variablen „*Wichtigkeit der Teamleistung*“ und „*Erfahrungsindex*“, verursacht durch die Variable „*Leistungsmotiv HE*“, kann ausgeschlossen werden.

Ebenso steht das *Residuum Positivität der sozio-emotionalen Aussagen* in einem höchst signifikanten positiven Zusammenhang mit dem *Residuum Witzanzahl* ( $r = .68, p < .001, 1 - \beta > .999$ ). Dies ist nicht verwunderlich, da die Witze einen Teil der positiven sozio-emotionalen Aussagen ausmachen. Interessanterweise wirken sich jedoch Witze stärker auf die Leistung der Entwickler aus als eine allgemein hin positive sozio-emotionale Interaktion.

Wie diese Ergebnisse aufzeigen, helfen Kontrollvariablen, die experimentell gefundenen Ausprägungen einer aV besser zu verstehen. Aufgrund der zu geringen Gesamtstichprobe in dieser Untersuchung konnten die vielfältigen Wechselwirkungen zwischen allen erhobenen Variablen statistisch nicht überprüft werden. Dennoch konnten durch Dodt (2012) und die Autorin einige praxisrelevante Zusammenhänge zwischen Input-, Prozessvariablen und der Leistung von Entwicklern von Automationen statistisch abgesichert oder zumindest eine entsprechende Tendenz nachgewiesen werden. Dadurch war es möglich, empirisch Faktoren

und Facetten des Faktors „Kommunikation intern“ für das „Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme inklusive Automationen“ mit der Output-Variable „Teamleistung“ (Rocholl, 2010) zu identifizieren und somit das Modell teilweise zu validieren respektive zu erweitern. In Abbildung 28 werden der Übersicht halber nur diese Faktoren im Rahmengerüst des Modells (ohne die Kategorie „Individuum“, da kein Faktor auf dieser Ebene betrachtet wurde) dargestellt.

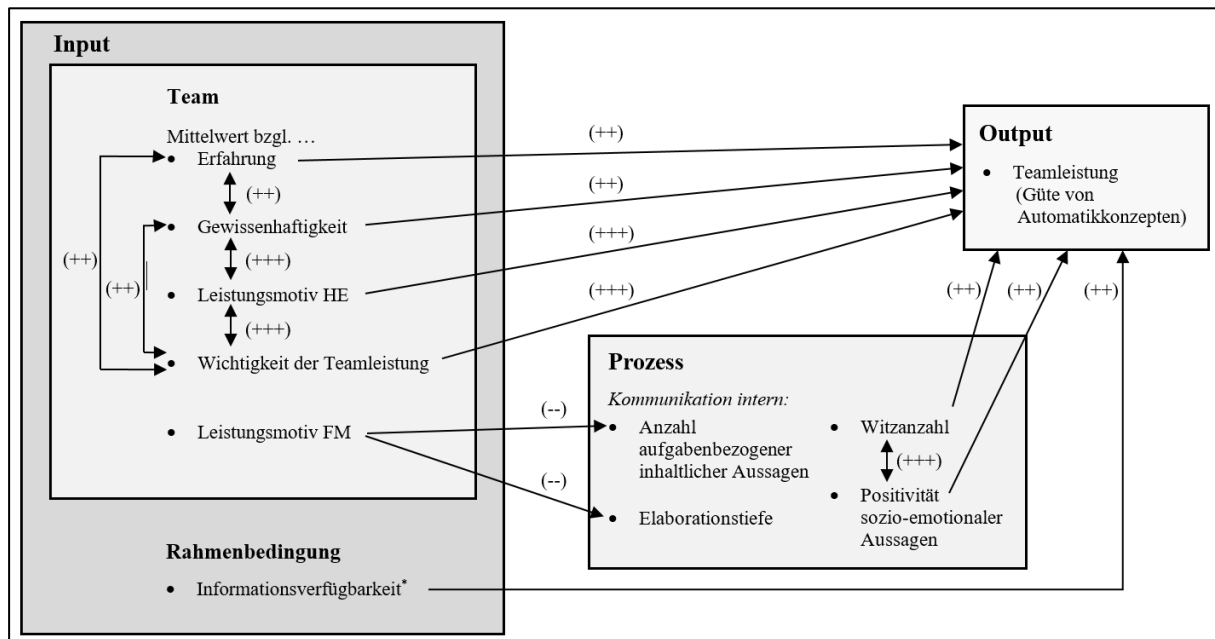


Abbildung 28: „Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme inklusive Automationen“ – Darstellung der empirisch gefundenen Zusammenhänge zwischen den Input-, Prozessvariablen und der Output-Variable

*Anmerkungen.* In dem Modell werden die Richtung (+ = positiv und - = negativ) und die Stärke (+/- = kleine, ++/-- = mittlere und +++/-- = große Effektgröße) eines Zusammenhangs einzelner Variablen dargestellt. Dabei ist zu beachten, dass die über die durchgehenden Linien dargestellten Kausalzusammenhänge nur aufgrund des zeitlichen Vorkommens der Variablen angenommen sind.; Für die Prozessvariablen und die Output-Variable „Teamleistung“ wurden bei der Bestimmung der Zusammenhänge Residuen verwendet, d. h. aus deren Varianz ist der durch die uV „Informationsverfügbarkeit“ aufklärbare Anteil herauspartialisiert. Einzig für den Zusammenhang mit der Input-Variable „Informationsverfügbarkeit“ wurde ein anderes Residuum der Teamleistung verwendet. Aus dessen Varianz ist der Anteil eliminiert, der durch die Variablen „Erfahrung“, „Leistungsmotiv HE“ und „Witzanzahl“ aufklärbar ist. Daher ist die Informationsverfügbarkeit mit einem \* gekennzeichnet.; Die bestimmten Rangkorrelationskoeffizienten tau von Kendall zwischen den Variablen „Wichtigkeit der Teamleistung“ als auch „Erfahrungsindex“ und dem „Residuum Güte der Automatikkonzepte“ wurden nach der Formel von Kendall (1979, zitiert nach Walker, 2003) in das Effektgrößenmaß r umgewandelt.

#### 5.2.3.4.2 Interpretation und Diskussion der Ergebnisse zur Forschungsfragestellung 5

##### 5.1 Forschungshypothesen bezüglich aV „Quantität der Automatikkonzepte“

Zwischen den drei Bedingungen der uV „Informationsverfügbarkeit“ zeigen sich hinsichtlich der, durch die Entwicklerteams konzipierten, Anzahl an Automaten, Anzahl an Funktionen und der durchschnittlichen Anzahl an Funktionen je Automat nach den Spezifika-



tionsbogen sowie der Anzahl an Funktionen nach der Bewertungsfunktionsmatrix keine signifikanten Unterschiede. Somit müssen die *Hypothesen F.5.1.1, F.5.1.2, F.5.1.3 und F.5.1.4*, in denen bzgl. dieser vier aVn Unterschiede zwischen den drei Bedingungen der uV „Informationsverfügbarkeit“ postuliert wurden, **verworfen** werden. Die, zur Überprüfung dieser Hypothesen ausgeführten, vier statistischen Analysen weisen nur kleine Teststärken auf. Für sichere statistische Überprüfungen sind größere Stichproben notwendig. Die Differenzen zwischen den drei Stichproben der Bedingungen hinsichtlich dieser vier aVn erreichen jedoch nicht die mittlere Effektgröße. Daher werden diese quantitativen Unterschiede der Automatikkonzepte zwischen den drei Bedingungen als vernachlässigbar erachtet.

### ***5.2 Forschungshypothese bezüglich aV „Qualität der Automatikkonzepte“***

Letztlich spielt es eine größere Rolle, ob die relevanten Schwachstellen im Prozess von SAM auf eine angemessene Art und Weise durch die konzipierten Automaten berücksichtigt werden. Bezüglich der qualitativen Differenzen der Automatikkonzepte zwischen den drei Bedingungen der uV „Informationsverfügbarkeit“ findet sich nur ein tendenziell signifikanter monotoner Trend in die erwartete Richtung. Dadurch muss die *Hypothese F.5.2* **verworfen** werden, nach welcher mit ansteigender Anzahl an Informationsquellen von der Bedingung 1 über die Bedingung 2 hin zur Bedingung 3 die Güte der Automatikkonzepte jeweils höher ausfällt. Jedoch weisen die Unterschiede zwischen den drei Bedingungen eine mittlere Effektgröße auf, die als praktisch bedeutsam gewertet wird. Die Wahrscheinlichkeit, ein signifikantes Ergebnis zu ermitteln, lag aufgrund der sehr geringen Teststärke lediglich bei 20 %. Somit ist zu konstatieren, dass für eine sichere statistische Überprüfung dieser Hypothese künftig größere Stichproben erhoben werden müssen.

### ***5.3 Kontrollhypothese bezüglich des moderierenden Einflusses der Kontrollvariablen auf die aV „Qualität der Automatikkonzepte“***

Um dennoch bereits im Rahmen dieser Untersuchung eine belastbarere Aussage über den experimentellen Effekt der uV „Informationsverfügbarkeit“ treffen zu können, wurde die interne Validität erhöht. Dafür wurden die Wirkungen der identifizierten relevanten Kontrollvariablen „Erfahrungsindex“, „Witzanzahl“ und „Leistungsmotiv HE“ auf die aV „Güte der Automatikkonzepte“ eliminiert. Dadurch konnte ein großer Teil der Störvarianz (57 %) in der aV aufgeklärt werden. Für die Wirküberprüfung der uV auf die aV war es somit möglich, die Teststärke zu erhöhen und einen höchst signifikanten, extrem praktisch bedeutsamen, hypothesenkonformen monotonen Trend zu ermitteln. Dabei zeigt sich, dass die uV von der verbliebenen Varianz der aV 37.20 % aufklären kann. Das wird hauptsächlich durch die signifikanten

und praktisch sehr bedeutsamen Unterschiede zwischen der Bedingung 1 und den anderen beiden Bedingungen verursacht. Zwischen den Bedingungen 2 und 3 findet sich zwar ein annähernd mittlerer Effekt, dieser kann aufgrund einer zu geringen Teststärke statistisch wiederum nicht abgesichert werden. Aufgrund dieser Ergebnisse kann die **Hypothese K.5.3**, dass Kontrollvariablen die Beziehung zwischen der uV „Informationsverfügbarkeit“ und der aV „Güte der Automatikkonzepte“ beeinflussen, **angenommen** werden. Die Ergebnisse untermauern zudem die Aussage von Bargh und Ferguson (2000), wonach menschliches Verhalten und dessen Leistung multipel determiniert sind.

Dennoch weist das Ergebnis bzgl. der qualitativen Differenzen der Automatikkonzepte zwischen den drei Bedingungen der uV „Informationsverfügbarkeit“ ohne Herauspartialisieren des Einflusses der Kontrollvariablen auf die aV die höhere ökologische Validität auf. In realen Entwicklungsprozessen besitzen die Mitglieder eines Teams unterschiedlichste Charakteristika, Fähigkeiten, Fertigkeiten, Erfahrungen und Wissen.

#### ***5.4 Explorative Analyse hinsichtlich der Unterschiede zwischen den Bedingungen der uV „Informationsverfügbarkeit“ bei den Kategorien der Bewertungsfunktionsmatrix***

Durch eine detaillierte explorative Analyse der Automatikkonzepte konnten Unterschiede zwischen den Bedingungen der uV „Informationsverfügbarkeit“ bzgl. der Kategorien der Funktionsmatrix gefunden werden. So berücksichtigten die Teams der Bedingung 1 im Vergleich zu den anderen beiden Bedingungen am häufigsten Funktionen zur Vermittlung von Informationen an die MWB. Eine der Einschränkungen der Konzeptionsaufgabe lautete, dass die MWB aktiver Bestandteil des Systems bleiben sollen und nur in Ausnahmefällen eine Automatik die Objektregelung zu 100 % übernehmen darf. Durch die Systembeschreibung als einzige Informationsquelle scheinen sich die Entwickler zur Berücksichtigung dieser Einschränkung auf die indirekte Objektregelung über die MWB konzentriert zu haben.

Die Teams der Bedingung 2 liegen hinsichtlich der Funktions- und Punktzahl bei der Unterkategorie „Entscheidung“ vor denen der anderen beiden Bedingungen. Möglicherweise identifizierten diese Teams als Hauptursache der Prozessstörungen von SAM die mangelnde Möglichkeit der MWB, sich gemeinsam für ein Regelungsverhalten zu entscheiden. Die Teams wurden eventuell durch das zuvor gezeigte Video von den Problemstellen beim kooperativen Tracking, das selbst Ausführen des kooperativen Trackings und das Interview mit den MWBn beeinflusst. Dadurch ist es möglich, dass diese Teams vor allem in der Automatisierung der erwähnten Entscheidung eine Optimierung des SAM-Prozesses antizipierten.

Die Teams der Bedingung 3 haben im Vergleich zu den anderen beiden Bedingungen am stärksten Funktionen für die Hauptkategorie „Handlungsausführung“ berücksichtigt. Dabei

konzipierten sie am häufigsten Funktionen zur adaptiven Veränderung des Objektregelungsanteils zwischen den MWBn sowie zur adaptiven Übernahme des Objektregelungsanteils durch die Funktion selbst. Gleichfalls planten diese Teams am häufigsten Funktionen zur graduellen Veränderung des Objektregelungsanteils zwischen den MWBn sowie zur graduellen Rückgabe der Objektregelung an die MWB. Vermutlich damit diese Funktionen adäquat agieren können, konzentrierten sich diese Teams ebenfalls am stärksten auf die Erfassung und Berechnung von objektregelungsrelevanten Daten. Scheinbar führten die verschiedenen Perspektiven auf SAM und den Operator zu einem zumindest tendenziell vielfältigeren Funktionsrepertoire. Vermutlich dienten die Informationsquellen „HAA bzgl. der Aufgaben des Operators“, „Screen-recording der Operateurstätigkeit am OA“ und „selbst als Operator agieren“ den Entwicklern als Ideenanker für diese Funktionalitäten der Automaten. So sind in der HAA (s. Anhang C.1.1) bereits die Kontrolle von relevanten Daten (bspw. Joystick-Bewegungen der MWB), die Ermittlung von Differenzen zwischen Ist- und Soll-Werten (bspw. Objektabweichung von Ideallinie) und die Antizipation von kritischen Ereignissen (bspw. Kollisionen mit Hindernissen) als notwendige Aufgaben zur Überwachung und Führung von SAM beschrieben. Sowohl in der HAA als auch im OA finden sich Möglichkeiten zur direkten Beeinflussung der Objektregelung (bspw. stufenweise Veränderung des Objektregelungsanteils zwischen den MWBn). Dabei kann der Operator den Grad der Beeinflussung variieren und somit an den Zustand der MWB oder Tracking-Situation anpassen. Dafür liegen dem Operator verschiedene Informationen vor, etwa zur Genauigkeit und Geschwindigkeit der Objektregelung. Erstaunlicherweise hat weder ein Team der Bedingung 3 noch eines der anderen beiden Bedingungen Funktionen zur Erfassung psychophysiologischer Daten der MWB konzipiert. Das verwundert insbesondere deshalb, da sowohl in der HAA als auch im OA die Beanspruchung der MWB als Systemzustandsinformation für den Operator vorgesehen ist.

Mit der HAA und dem OA kann darüber hinaus begründet werden, warum in Bedingung 3 am häufigsten objektregelungsachsenspezifische Funktionen geplant wurden. Beide Informationsquellen weisen solche Funktionalitäten auf. Hierzu zählen z. B. die Limitierung der vertikalen Objektgeschwindigkeit oder die Blockierung einer Seite der horizontalen Achse.

Des Weiteren kann der OA dazu geführt haben, dass die Teams der Bedingung 3 keine Funktionen zur Anzeige von Ist-Werten konzipierten. So bietet der OA als eine Systemzustandsinformationsquelle dem Operator die Ist-Joystick-Eingaben der MWB in Echtzeit an. Bei der Fülle an Elementen auf dem GUI des OAs und der Dynamik der zu überwachenden Tracking-Aufgabe, könnte diese Information für die Entwickler nur wenig Mehrwert zur Ausführung der Aufgabe geboten haben, während sie selbst als Operator fungierten. Diese

Erfahrung wurde möglicherweise auf die ebenfalls dynamische Situation der MWB übertragen. Daher könnten Funktionen zur Vermittlung von Ist-Wert-Informationen im Vergleich zu anderen Funktionalitäten als weniger relevant eingestuft worden sein.

Die Pfadvorausschau im OA dürfte der Grund sein, weshalb nur Teams der Bedingung 3 diese Funktionalität für die MWB berücksichtigten.

Außerdem könnte im Agieren der Entwickler als Operateur in der Bedingung 3 die Ursache dafür liegen, weshalb sie im Vergleich mit den Teams der anderen Bedingungen am seltensten Funktionen konzipierten, welche die MWB individuell berücksichtigen. Die individuelle Berücksichtigung eines MWBs über den OA erfordert durch eine Zweischnittbedienung den höchsten Bedienaufwand. Möglicherweise haben die Entwickler aufgrund der Dynamik der Tracking-Aufgabe auf diesen Aufwand im Verlauf der Zeit verzichtet und konnten dabei keinen Unterschied im Objektregelungsverhalten der MWB feststellen. Zudem könnte es sein, dass die Entwickler beim Agieren als MWB in den individuellen Anweisungen des Operateurs keinen Mehrwert wahrnahmen.

Nahezu die Hälfte der gefundenen Unterschiede zwischen den drei Bedingungen der uV „Informationsverfügbarkeit“ hinsichtlich der Kategorien der Funktionsmatrix weisen eine mittlere und einige sogar eine hohe Effektgröße auf (s. Anhang G.6.6). Jedoch konnten diese, aufgrund zu geringer Teststärken ( $1 - \beta = .658$ ) der angewendeten Analyseverfahren, statistisch nicht abgesichert werden. Trotzdem führen diese Ergebnisse zu der Annahme, dass die, den Entwicklern zur Verfügung stehenden, Informationsquellen den Rahmen für den Lösungsraum bilden, aus dem die Funktionalitäten zur Lösung einer Aufgabenstellung entnommen werden. Daher ist es ratsam, für eine Entwicklungstätigkeit Informationsquellen heranzuziehen, welche die unterschiedlichsten Perspektiven auf ein Problem bieten, um so den Lösungsraum angemessen zu gestalten. Dabei sollten auch ähnliche Produkte von Wettbewerbern mit einbezogen werden (J. Schwarz, 2007). Wobei Informationsquellen, welche potentielle Lösungsideen beinhalten, nicht zum Entwicklungsbeginn eingesetzt werden sollten. Denn nur so kann kreativen, innovativen Lösungsideen eine Chance zur Entfaltung gegeben werden.

### ***5.5 Erklärungsansätze für die gefundenen Ergebnisse zur Forschungsfragestellung 5***

Einen möglichen Grund für den nur annähernd mittleren und tendenziell signifikanten Effekt der systematisch variierten uV „Informationsverfügbarkeit“ im Verbund mit den Kontrollvariablen auf die aV „Güte der Automatikkonzepte“ erwähnte bereits Krinner (2009). Sie nahm an, dass das, von ihr in ihren Untersuchungen verwendete, vereinfachte Lastenheft in dem das System beschrieben wurde, aufgrund einer zu detaillierten Gestaltung einen Deckeneffekt verursachte. Dadurch konnten die anderen Informationsquellen keinen signifikanten

inhaltlichen Mehrwert leisten. Die, im Rahmen dieser Untersuchung verwendete, Systembeschreibung wurde neu angefertigt und inhaltlich auf das Wesentliche zur Beschreibung der Logik von SAM reduziert. Dennoch ist es möglich, dass durch diese Systembeschreibung bereits ein Großteil der entwicklungsrelevanten Informationen an die Entwickler vermittelt wurde. Diefenbach et al. (2013) fanden heraus, dass ein Verständnis von der Logik eines Konzepts mittels einer textuellen Beschreibung, die sie u. a. mit einem Video verglichen, am besten aufgebaut werden kann. Ebenso argumentieren sie, dass bei einem additiven Design, wie es in dieser Untersuchung verwendet wurde, durch die Basisquelle, in diesem Fall die Systembeschreibung, die spezifischen Effekte der hinzukommenden Informationsquellen verdeckt werden können. Infolgedessen fällt der Unterschied zwischen den Stufen der uV vermindert aus.

Die Annahme eines Deckeneffekts der Systembeschreibung, der die Wirkung der uV auf die aV abgeschwächt hat, wird durch die Beurteilungen der Entwickler hinsichtlich der Wichtigkeit der einzelnen Informationsquellen für die Konzeptionsaufgabe verstärkt. So ergibt sich aus der gebildeten Rangreihe der Informationsquellen nach deren Wichtigkeit unter allen drei Bedingungen für die Systembeschreibung der durchschnittlich höchste Rangplatz. Wobei interessanterweise der durchschnittliche Rangplatzwert unter Bedingung 2 ( $M_{\text{TeamMRangB2}} = 1.22$ ,  $SD_{\text{TeamMRangB2}} = 0.28$ ) gefolgt von Bedingung 3 ( $M_{\text{TeamMRangB3}} = 1.83$ ,  $SD_{\text{TeamMRangB3}} = 1.11$ ) am höchsten ausfällt ( $M_{\text{TeamMRangB1}} = 2.93$ ,  $SD_{\text{TeamMRangB1}} = 2.68$ ).<sup>138</sup> Scheinbar wird die Systembeschreibung im erlebten Kontext der anderen Informationsquellen als noch wichtiger empfunden. Zudem wurde die Systembeschreibung als annähernd völlig wichtig bewertet ( $M_{B1} = 4.60$ ,  $SD_{B1} = 0.89$ ,  $M_{B2} = 4.83$ ,  $SD_{B2} = 0.46$ ,  $M_{B3} = 4.63$ ,  $SD_{B3} = 0.49$ )<sup>139</sup>. Dabei erhielt sie zugleich den höchsten Wichtigkeitswert.

Wie den Beurteilungen der Wichtigkeit der Informationsquellen durch die Entwickler zudem entnommen werden kann (s. Anhang G.6.4), wurden in der Bedingung 3 die hier erstmalig vorhandenen Informationsquellen (Perspektive auf den Operateur) als am unbedeutendsten wahrgenommen. Eine Ausnahme stellt die Informationsquelle „selbst als Operateur agieren“ dar, welche bei der Einschätzung der Höhe der Wichtigkeit den dritthöchsten Wert erreichte ( $M_{B3} = 3.5$ ,  $SD_{B3} = 1.00$ ). Eine zweite Ausnahme stellt die Informationsquelle „MWB

---

<sup>138</sup> Dieser globale Unterschied zwischen den Bedingungen weist einen annähernd großen Effekt auf, welcher aufgrund der geringen Teststärke des Analyseverfahrens statistisch nicht abgesichert werden konnte ( $H[2] = 2.385$ ,  $p = .303$ ,  $\eta^2_p = .159$ ,  $1 - \beta = .494$ ).

<sup>139</sup> Bei den Beurteilungen der Wichtigkeit der Informationsquellen werden nicht die durchschnittlichen Team-Mittelwerte, sondern die durchschnittlichen individuellen Entwicklereinschätzungen berichtet. Wegen der Antwortoption „weiß nicht“ war es bei zu vielen Teams unmöglich, einen Mittelwert zu bilden.

interviewen“ dar, welche bei der Rangreihe der Wichtigkeit der Informationsquellen nur auf Platz 6 ( $M_{\text{TeamMRangB3}} = 5.63$ ,  $SD_{\text{TeamMRangB3}} = 1.65$ ) hinter den Informationsquellen „selbst als Operateur agieren“ (Platz 4,  $M_{\text{TeamMRangB3}} = 4.58$ ,  $SD_{\text{TeamMRangB3}} = 1.45$ ) und „Screenrecording der Operateurstätigkeit am OA“ (Platz 5,  $M_{\text{TeamMRangB3}} = 5.35$ ,  $SD_{\text{TeamMRangB3}} = 0.84$ ) liegt. Diese Wichtigkeitsbeurteilungen der Informationsquellen helfen nachzuvollziehen, warum der Unterschied zwischen Bedingung 2 und 3 hinsichtlich der aV „Güte der Automatikkonzepte“ selbst nach der Eliminierung der Wirkung der Kontrollvariablen am geringsten ausfällt.<sup>140</sup> Es könnte sein, dass das zusätzliche Wissen zum Operateur nur wenig Mehrwert bot. Denn schließlich sollten Automaten für SAM konzipiert werden, welches ein autonomes System ohne Operateur ist.

Des Weiteren wird die Annahme eines Deckeneffekts der Systembeschreibung durch die Einschätzung der Entwickler gestützt, wie ausreichend die zur Verfügung gestellten Informationsquellen für die Konzeptionstätigkeit waren. Zwar nimmt mit steigender Anzahl an Informationsquellen im Durchschnitt die Einschätzung zu, dass diese ausreichend waren. Dennoch zeigt sich erst ein globaler mittlerer Unterschied zwischen den Bedingungen nach der Eliminierung der zwei Ausreißer der Bedingung 3, welcher aufgrund der geringen Teststärke statistisch nicht abgesichert werden konnte ( $M_{\text{TeamMB1}} = 3.80$ ,  $SD_{\text{TeamMB1}} = 0.65$ ,  $M_{\text{TeamMB2}} = 3.93$ ,  $SD_{\text{TeamMB2}} = 0.61$ ,  $M_{\text{TeamMB3}} = 4.17$ ,  $SD_{\text{TeamMB3}} = 0.18$ ,  $H[2] = 2.331$ ,  $p = .312$ ,  $\eta^2_p = .076$ ,  $1 - \beta = .230$ ). Ebenso werden die Informationsquellen in allen Bedingungen als überwiegend ausreichend für die Konzeptionsaufgabe eingeschätzt.

Aufgrund des angenommenen Deckeneffekts der Systembeschreibung, der Beurteilungen der Informationsquellen hinsichtlich ihrer Wichtigkeit und wie ausreichend diese sind, sollte in zukünftigen Untersuchungen auf ein additives Design verzichtet werden, wenn die Wirkungen der einzelnen Informationsquellen im Forschungsfokus stehen. Dass die anderen Informationsquellen ebenfalls eine Rolle spielen, zeigt sich in den Angaben der Entwickler zu den Verständnisproblemen während der Untersuchung. So wurde bspw. angegeben: „*Dokumentation war sicher sachlich richtig, vollständig und vermutlich ohne Definitionslücken, kann aber nicht das Ausprobieren bzw. Beobachten ersetzen.*“

Erwähnt werden muss auch, dass sich die Einschätzung der Wichtigkeit der beiden Informationsquellen „MWB interviewen“ und „Operateur interviewen“ verringerte, sobald diese tatsächlich zur Verfügung standen. So liegt die Informationsquelle „MWB interviewen“

---

<sup>140</sup> Aufgrund einer zu geringen Teststärke des Analyseverfahrens konnte dieser annähernd mittlere Unterschied statistisch nicht abgesichert werden (s. Anhang G.6.5).

hinsichtlich der Wichtigkeitseinstufung in Bedingung 1 auf Platz drei ( $M_{B1} = 3.67$ ,  $SD_{B1} = 1.18$ ), in Bedingung 2 auf Platz sechs ( $M_{B2} = 3.03$ ,  $SD_{B2} = 1.09$ ) und in Bedingung 3 auf Platz fünf ( $M_{B3} = 3.23$ ,  $SD_{B3} = 0.68$ ). Die Informationsquelle „Operator interviewen“ befindet sich in Bedingung 1 noch auf Rangplatz vier ( $M_{\text{TeamMRangB1}} = 5.48$ ,  $SD_{\text{TeamMRangB1}} = 1.44$ ), fällt in Bedingung 2 auf Platz neun ( $M_{\text{TeamMRangB2}} = 7.65$ ,  $SD_{\text{TeamMRangB2}} = 0.85$ ) und in Bedingung 3 auf Platz sieben ( $M_{\text{TeamMRangB3}} = 6.12$ ,  $SD_{\text{TeamMRangB3}} = 1.15$ ) ab. Die schlechten Bewertungen in Bedingung 2 können durch die Übertragung der Erfahrung mit dem Interview der MWB verursacht worden sein. Vermutlich wurden beide Interviews als weniger wichtig eingestuft, wenn diese erlebt wurden, weil keine echten MWB oder Operateure befragt werden konnten. Diese Annahme wird bekräftigt durch die Angabe eines Entwicklers der Bedingung 3 auf die Frage, welche Informationsquellen gefehlt hätten: „*echte Interviews mit betroffenen Personen*“. Zudem wurden die Antworten von Krinner (2009) und der Autorin generiert. Dabei wurde darauf geachtet, die Entwickler bei der Lösungsfindung nicht in eine bestimmte Richtung zu beeinflussen. Daher wurden alle relevanten Schwachstellen von SAM passend zu der jeweiligen Frage dargestellt. Reale MWB und Operateure hätten vermutlich Schwerpunkte bei der Problemdarstellung hinsichtlich der Relevanz für die eigene Aufgabe gesetzt. Ferner ist anzunehmen, dass diese auch bereits erste Lösungsideen mitgeliefert hätten (s. Sarodnick & Brau, 2006). Wenn in zukünftigen Untersuchungen erneut Interviews eingesetzt werden und reale MWB oder Operateure dafür nicht in Echtzeit zur Verfügung stehen, sollten zumindest die Antworten in den beiden Frage-Antwort-Katalogen von ihnen generiert worden sein.

Ferner liefert Krinner (2009) drei weitere Gründe für eine geringe Auswirkung einer uV auf die aV unter Einsatz von SAM, welche ebenfalls für die gefundenen Effekte in dieser Untersuchung mit verantwortlich sein können. So könnte zum einen ein Ankereffekt (Tversky & Kahneman, 1974, zitiert nach Krinner, 2009) aufgetreten sein. Danach kann ein bereits aus dem Alltag bekanntes System die Basis für die Beurteilung der eigenen Lösung darstellen. Ebenso könnte ein bestimmter kognitiver Frame (Kahneman & Tversky, 1984; Tversky & Kahneman, 1981, zitiert nach Krinner, 2009) bestanden haben. Bereits bekannte Systeme setzten den Rahmen, in dem sich die Lösungen befinden können. Ferner könnte das Prinzip des Satisficing (Gigerenzer, Todd, & ABC Research Group, 1999; Simon, 1956, 1995, zitiert nach Krinner, 2009) verwendet worden sein. Demzufolge beenden Entwickler die Suche nach alternativen Lösungen, sobald eine Lösung ein angemessenes Aspirationsniveau erreicht hat, welches wiederum durch ein bereits bekanntes System definiert wird. Wie bei den Untersuchungen von Krinner (2009) könnte es sich bei den, zur Konzeptionstätigkeit herangezogenen, Systemen aus dem Alltag um Fahrerassistenzsysteme oder Rennsimulationen handeln. Jedes Team nahm

während der Untersuchung Bezug zu realen Fahrerassistenzsystemen und einige der Teams zudem auf Rennspiele. Dies erfolgte trotz der Bemühungen in der Beschreibung von SAM und den Instruktionen, Begrifflichkeiten zu vermeiden, welche im Zusammenhang mit einer Fahrzeugführung stehen. Ziel war es, dass SAM prototypisch für komplexe, dynamische, technische Systeme wahrgenommen wird. Es sollte jedoch kein spezielles System darstellen. Allerdings ist der visuelle Eindruck von einer rudimentären Fahrsimulation in SAM zu stark ausgeprägt, wodurch sowohl Assoziationen auf Fahrsituationen und bereits existierende Fahrerassistenzsysteme wie auch auf Rennsimulationen nicht vermieden werden konnten.

### 5.2.3.5 Kritik

Es sind jedoch auch einige Aspekte an der Untersuchung zu bemängeln. Als ein kritischer Aspekt dieser Untersuchung könnte die Konzeptionsaufgabe selbst gesehen werden. Trotz der Bemühungen der Autorin, die Komplexitätswirkung von SAM auf Entwickler zu erhöhen (s. Kapitel 3), ist nicht auszuschließen, dass die Konzeption von Automaten für dieses System für erfahrene Entwickler zu einfach ist. Dadurch wäre ein Entwicklerteam unabhängig von der zur Verfügung stehenden Information in der Lage, alle Schwachstellen von SAM zu erkennen und angemessene Automaten zu konzipieren. Allerdings wurde die Komplexität der Konzeptionsaufgabe von den Entwicklern der drei Bedingungen im Durchschnitt an der Grenze zwischen „mittelmäßig bis überwiegend komplex“ eingeschätzt ( $M_{\text{TeamMB1}} = 10.57$ ,  $SD_{\text{TeamMB1}} = 18.53$ ,  $M_{\text{TeamMB2}} = 11.67$ ,  $SD_{\text{TeamMB2}} = 10.25$ ,  $M_{\text{TeamMB3}} = 9.87$ ,  $SD_{\text{TeamMB3}} = 16.85$ ,  $F[2, 27] = 0.034$ ,  $p = .967$ ,  $\eta^2_p = .002$ ,  $1 - \beta = .057$ ). Zugleich stuften die Entwickler in allen drei Bedingungen ihre Beanspruchung durch die Konzeptionstätigkeit im Mittel auf „eher viel Anstrengung“ ein ( $M_{\text{TeamMB1}} = 55.83$ ,  $SD_{\text{TeamMB1}} = 18.26$ ,  $M_{\text{TeamMB2}} = 55.43$ ,  $SD_{\text{TeamMB2}} = 11.44$ ,  $M_{\text{TeamMB3}} = 55.13$ ,  $SD_{\text{TeamMB3}} = 8.61$ ,  $H[2] = 0.072$ ,  $p = .965$ ,  $\eta^2_p < .001$ ,  $1 - \beta = .051$ ). Ebenso **sank** die **Stimmung** über die **Zeit** (Haupteffekt für Prä-Post:  $F[1, 27] = 4.573$ ,  $p = .042$ ,  $\eta^2_p = .145$ ,  $1 - \beta = .992$ ), was durch eine beanspruchende Aufgabe verursacht werden kann (s. Kaufmann et al., 1982). Außerdem wurde in 13 Teams von jeweils mindestens einem Entwickler zu wenig zur Verfügung stehende Zeit für die Konzeptionsaufgabe bemängelt (s. Anhang G.6.3). So äußerte ein Entwickler bei der Verabschiedung: „*Man müsste sich eigentlich mehrere Wochen damit beschäftigen.*“ Somit scheint die Konzeptionsaufgabe aufgrund der zur Verfügung stehenden Zeit für die Entwickler ausreichend herausfordernd gewesen zu sein. Eine nachfolgende Untersuchung untermauert diese Annahme. Erfahrene Studententeams in der Softwareentwicklung konzipierten lediglich für die Gabelungen von SAM, trotz dreiwöchiger Bearbeitungszeit, mit durchschnittlich etwas mehr als 21 Stunden Besprechungszeit im Team,



gleichwertige Automatikkonzepte (Niestroj, 2014). Bei einigen der Teams konnte aufgrund von zusätzlicher Bearbeitungszeit eine Qualitätssteigerung der Konzepte verzeichnet werden. Daher könnte die geringe Wirkungsentfaltung der uV „Informationsverfügbarkeit“ genauso gut an einem zu kurzen Zeitraum für die Konzeptionstätigkeit gelegen haben. Die verwendete Zeit zur Konzeption eines Produkts steht in einem positiven Zusammenhang mit dessen Güte (Hacker et al., 2002). Eine Konzeptionsarbeit innerhalb von zwei Stunden ist weit entfernt von der zur Verfügung stehenden Zeit bei realen Entwicklungstätigkeiten.

Ferner könnte das Untersuchungssetting nicht geeignet gewesen sein, um weitere, in der Entwicklungspraxis vorherrschende, bedeutsame Zusammenhänge aufzudecken. Andere Untersuchungen belegen, dass sich bestimmte Zusammenhänge nur in Feldstudien zeigen (Bell, 2004, 2007). Die hier vorgestellte Untersuchung ist jedoch auf dem Kontinuum zwischen den Extremen „Laboruntersuchung“ und „Feldstudie“ angesiedelt. Durch den Einsatz von professionellen Entwicklern in ihrem natürlichen Arbeitsumfeld sollten Eigenschaften einer Feldstudie erreicht werden. Aufgrund der künstlichen und kontrollierten Konzeptionsaufgabe weist die Untersuchung jedoch auch Eigenschaften einer Laboruntersuchung auf, welche die Ergebnisse auf spezifische Weise beeinflusst haben könnten. So merkte bereits Dodt (2012) an, dass die im Verlauf der Untersuchung vollbrachte Leistung der Entwickler, anders als bei realen Entwicklungstätigkeiten, für ihr weiteres Leben keine Bedeutung hat. Die Entwickler hatten bei schlechter Leistung keine Sanktion zu befürchten. Zugleich sollte nur das beste Team mit einem Amazon-Gutschein i. H. v. 50 Euro belohnt werden. Daher kann die aufgebrachte Motivation der Entwickler, zur Erbringung einer guten Leistung, geringer ausgefallen sein, als dies in realen Entwicklungstätigkeiten der Fall ist. Damit ginge eine verminderte ökologische Validität der Untersuchung einher. Bereits in anderen Untersuchungen konnten positive Zusammenhänge zwischen der Motivation und dem Bemühen, ein Ziel zu erreichen (R. H. Rasch & Tosi, 1992), sowie der erwarteten Belohnung und der erreichten Leistung (Levi & Slem, 1995; Susman & Ray, 1999) nachgewiesen werden.

Ebenso mussten sich die Entwickler kaum einschränkende Gedanken hinsichtlich der Umsetzbarkeit der Automatikkonzepte machen, da sie diese nicht zu implementieren hatten. Anders als in den Untersuchungen von Krinner (2009) wünschte sich keiner der Entwickler, die Implementierung der Automaten vornehmen zu können. Es fielen Aussagen, wie: „*Zum Glück mussten wir das nicht umsetzen.*“. Dieses Wissen um diese Arbeitsteilung spiegelt sich in den teilweise rudimentär spezifizierten Automatikkonzepten wider.

Eine erste Betrachtung, welche Auswirkungen eine Aufgabe, bestehend aus Konzeption und Implementierung, auf die Leistung von Entwicklern hat, wurde bereits im Anschluss an

diese Arbeit durch Niestroj (2014) vorgenommen. Er konnte keinen Güteunterschied zwischen nur konzipierten und zusätzlich implementierten Automatikkonzepten ausfindig machen. Allerdings wurde diese Untersuchung mit Studenten durchgeführt, die für die Erfüllung der Aufgabe Studienpunkte erhielten. Daher mögen sie eventuell höher motiviert gewesen sein, praktisch umsetzbare Konzepte zu erstellen, als es die Entwickler in der hier vorgestellten Untersuchung waren.

Dodt (2012) erwähnte ebenfalls bereits, dass die Kameraaufnahmen das Verhalten der Entwickler ergebnisbeeinflussend verändert haben könnten. Bei keinem der Teams konnte die Kamera gänzlich unauffällig positioniert werden. Die Mikrofone lagen zudem eher auffällig vor jedem Team. Diese Präsenz der Aufnahmegeräte spiegelte sich in den, von allen Teams getätigten, Scherzen über die Aufnahmen wider. Um ökologisch valide Daten der Entwicklungstätigkeit zu erheben, sollten die Entwickler so wenig wie möglich gestört werden (Hacker & Sachse, 2006).

Abschließend ist der Erhebungszeitpunkt der Input-Variablen als kritisch zu erwähnen. So wurden diese erst nach der Konzeptionsphase erfragt. Der Prozess und das Ergebnis können sich somit auf einige dieser Variablen ausgewirkt haben. Selbst die uV kann einige dieser Variablen beeinflusst haben, wie etwa die „Wichtigkeit der Teamleistung“. So steigert eine als angemessen empfundene Informationsmenge die Motivation et vice versa (Spieß & Winterstein, 1999, zitiert nach Dodt, 2012). In zukünftigen Untersuchungen sollten diese Variablen zum Beginn oder durch Prä-Post-Messungen erhoben werden, um Beeinflussungen zu vermeiden bzw. sogar identifizieren zu können.

Trotz der zuvor genannten kritischen Aspekte, ist es gelungen, die Bedeutung der Leistungsdeterminante „Informationsverfügbarkeit“ für die Güte von Automatikkonzepten aufzudecken. Vergleichend mit den Untersuchungen von Krinner (2009) konnte dies durch den Einsatz einer größeren Anzahl an Informationsquellen, die unterschiedlichen Perspektiven auf SAM hatten, und durch die Teilnahme professioneller Entwickler an der Untersuchung sowie die detailliertere Analyse der Automatikkonzepte unter Einbeziehung von Kontrollvariablen erreicht werden.

#### **5.2.3.6 Ausblick**

Im Rahmen dieser Dissertation konnten aus zeitökonomischen Gründen einige für eine Entwicklungstätigkeit relevante Aspekte nicht analysiert werden. So könnte bspw. anhand der Daten dieser Untersuchung überprüft werden, ob die professionellen Entwickler tatsächlich strukturierter vorgehen, als es die Studenten (s. Ahmed et al., 2003; Anderson, 2001; Badke-Schaub, 2005; Christiaans & Dorst, 1992; G. Fricke, 1993) in den Untersuchungen von Krinner

(2009) und Niestroj (2014) taten. Zugleich wäre interessant, ob sich die uV „Informationsverfügbarkeit“ vermittelt über das Vorgehen auf die Güte der Automatikkonzepte ausgewirkt hat. Beobachtet werden konnte, dass die meisten Teams, zumeist in der ersten halben Stunde, viel Zeit damit verbrachten, die Aufgabe und ihr Team zu strukturieren und strategisch zu organisieren. Dabei zeigten sich teilweise stark wechselhafte Themeninhalte. Im Laufe der Konzeption fanden zudem immer wieder Neuorganisationen statt. Dies stimmt mit Befunden von Badke-Schaub und Stempfle (2003) überein, nach denen fast 1/3 der Bearbeitungszeit einer Aufgabe in die Strukturierung und Organisation der Gruppe investiert wird. Solch ein Vorgehen ist für die Bewältigung von komplexen Entscheidungsaufgaben im Team (Kolbe, 2007) zur Erreichung einer hohen Teamleistung nötig (Scholl, 2005). Ebenso vollführten alle Teams innerhalb der zweiten Konzeptionsstunde einen zumeist abrupten Strategiewechsel. Von einer zunächst vorrangig verbalen Teamkonzeptionstätigkeit gingen sie zum systematischen Ausfüllen der Spezifikationsbogen über, wobei dies oftmals in Einzelarbeit geschah. Es ist anzunehmen, dass ein entstandenes Zeitdruckempfinden die Entwickler zu diesem Strategiewechsel veranlasste, um so die Konzeptionstätigkeit angemessen abschließen zu können (s. Frankenberger, 1997).

Ferner könnte anhand der erfassten Daten beleuchtet werden, ob die erfahrenen Entwickler ein Cognitive Tunneling (s. Dirkin, 1983) zeigten und dies unter einer der Bedingungen der uV „Informationsverfügbarkeit“ verstärkt vorkam. Denn in den Untersuchungen von Krinner (2009) neigten Studenten mit einem mittleren Effekt vorrangig dazu, wenn ihnen nur das vereinfachte Lastenheft als Informationsquelle zur Verfügung stand.

Darüber hinaus könnte anhand der schriftlichen, individuellen sowie der mündlichen, gemeinsamen System- und Konfliktbeschreibung durch die Entwickler, welche sie direkt im Anschluss an die Phase „Exploration der Informationsquellen“ tätigten, überprüft werden, inwiefern die Entwickler in den drei Bedingungen ein unterschiedliches Systemverständnis aufbauten oder unterschiedliche Konfliktstellen identifizierten. Ebenso könnten die individuellen Angaben mit denen des jeweiligen Teams verglichen werden. Dabei wäre bspw. die Frage zu klären, ob es den einzelnen Entwicklern gelang, ihr jeweils gewonnenes Einzelwissen in einem Teamwissen zu integrieren. Die Ausbildung eines gemeinsamen mentalen Modells der Problemstellung erleichtert eine koordinierte Zusammenarbeit (Wright & Kaber, 2005) und trägt zum Erfolg einer Konzeptionstätigkeit bei (Tiwana, 2004, zitiert nach Rocholl, 2010).

## 6. Gesamtfazit und Ausblick

Insgesamt stellt diese Arbeit eine Zwischenstufe auf dem Weg zur Optimierung der Funktionsteilung zwischen Mensch und Maschine dar. Einige interessante und praxistaugliche Erkenntnisse konnten gewonnen werden. Nachfolgend soll noch einmal auf wesentliche Aspekte dieser Arbeit eingegangen und dazu ein Ausblick gegeben werden.

### *1. Eignung der Versuchsumgebung SAM*

In der Versuchsumgebung SAM sollen die Prozesse von interagierenden Variablen realer komplexer, dynamischer, technischer Systeme durch die dynamische Interaktion der beiden MWB dargestellt werden (Gross, 2008). Die dynamische Interaktion der MWB sollte zu einer von den Entwicklern und dem Operateur wahrnehmbaren Komplexität von SAM führen, durch welche die Überwachungs- und Führungsaufgabe von SAM für die Entwickler über deren Automaten und für den Operateur herausfordernd und vergleichbar mit deren realen Arbeitstätigkeiten ist. Für professionelle Entwickler scheint die Konzeption von Automaten zur Überwachung und Führung des SAM-Prozesses in einem Zeitrahmen von zwei Stunden herausfordernd zu sein, da diese in der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“ (s. Abschnitt 5.2.3) eine solche Aufgabe als eher komplex und anstrengend beurteilten. Zudem kann, aufgrund der erreichten Variation der Automatenkonzepte in dieser Untersuchung, davon ausgegangen werden SAM prinzipiell für die experimentelle Erforschung der Entwicklungsprozesse von Automationen geeignet ist.

Eine Ermittlung der tatsächlich erreichten Höhe der ökologischen Validität von SAM steht noch aus. Dafür werden Kriterien zur Bestimmung der ökologischen Validität einer solchen Versuchsumgebung benötigt. Hierbei kann sich an der Arbeit von Gonzalez et al. (2005) orientiert werden, welche Mikrowelten hinsichtlich deren Undurchsichtigkeit, Dynamik, Komplexität sowie dynamischen Komplexität analysierten.

Eine weitere Zugangsmöglichkeit zur Ermittlung der ökologischen Validität stellt eine Aufgabenanalyse dar, bspw. die HAA (s. Abschnitt 3.3.1.1; Annett & Duncan, 1967). Mittels dieser können Vergleiche zwischen Systemen hinsichtlich der auszuführenden Aufgaben vorgenommen werden. Für solche Vergleiche sollten reale unterschiedlich komplexe und dynamische Systeme herangezogen werden, in denen ebenfalls Automationen und Operateure agieren. (Nachtwei & Kain, 2010)

Eine Möglichkeit, dass SAM nicht mehr als rudimentärer Fahrsimulator wahrgenommen wird, wäre bspw. das Layout der Tracking-Oberfläche zu verändern. Hierbei könnten andere Formen, Farben oder gar Muster für einzelne Bereiche verwendet werden. Dabei sollte

darauf geachtet werden, dass das Layout für die MWB über einen längeren Zeitraum ohne Anstrengung betrachtet werden kann. Bereits das kräftige Grün des Randbereichs des Tracking-Pfads wurde von einigen als MWB agierenden VPn als unangenehm für eine längere Betrachtung bemängelt. Ebenso sollten die verwendeten Formen und Farben nicht zu salient gestaltet werden, damit visuelle Anzeigen durch Operateure oder Automaten weiterhin wirken können.

## ***2. Identifizierung der Leistungsdeterminanten von Entwicklern***

Nach wie vor liegt nur spärliches praxisrelevantes Wissen über die vielfältigen Leistungsdeterminanten der Entwickler von komplexen, dynamischen, technischen Systemen inklusive Automationen und deren Wirkzusammenhänge vor. Ein Hauptziel des ATEO-Projekts war es, diese Leistungsdeterminanten zu identifizieren. Diese Dissertation konnte einen kleinen Wissensbeitrag dazu leisten. Jedoch sind weitere, umfangreichere Forschungstätigkeiten notwendig, um das komplexe Themenfeld aufklären zu können. Ein Wunsch der von Peeters et al. (2008) geteilt wird. Verschiedenste Output-Variablen und deren Rückwirkung auf die Input- und Prozessvariablen wie auch die Bedeutung dieser Wirkfaktoren für die einzelnen Entwicklungsphasen sollten betrachtet werden. So konnten bereits Maidique und Zirger (1985) aufzeigen, dass eine schlechte Leistung aufgrund von Lernprozessen die Wahrscheinlichkeit eines Erfolgs im nachfolgenden Projekt erhöht. Ebenso sollten Variablen respektive deren Ausprägungen mit negativen Einflüssen auf die Leistung von Entwicklern näher beleuchtet werden, wie irrelevante Informationsquellen (s. Diefenbach et al., 2013; Wübbelmann, 2001), um deren Wirkung in der Praxis vermindern oder gar eliminieren zu können. Für diese Forschungstätigkeiten bietet das, im Rahmen dieser Dissertation entstandene, „Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme“ eine gute Ausgangsbasis. Diesem kann entnommen werden, welche Variablen als potentielle Leistungsdeterminanten in Untersuchungen mit Entwicklern von Automationen systematisch überprüft, kontrolliert oder ausgeschaltet werden sollten. Dadurch kann nach und nach ein experimentell validiertes „Modell der Leistungsdeterminanten von Entwicklern komplexer, dynamischer, technischer Systeme mit Automationen“ entstehen. Die Wirkung einiger Faktoren konnte mit dieser Arbeit bereits belegt werden. Je besser das Modell validiert ist, desto gewinnbringender ist es für den Einsatz in der Praxis. So könnten zukünftig darauf basierend, aufgrund der darin gelisteten Erkenntnisse, Entwicklerteams zusammengestellt sowie deren Aufgaben, Rahmenbedingungen und Arbeitsprozess gestaltet werden. Dadurch sollten bessere Entwicklungsergebnisse erreicht werden können, z. B. eine optimalere Mensch-Maschine-Funktionsteilung.

### ***3. Einbezug von Kontrollvariablen***

Diese Arbeit konnte aufzeigen, dass durch das Einbeziehen von Kontrollvariablen in die inferenzstatistische Analyse einer uV, Aussagen zu deren Wirkung sicherer getätigt werden können. Dieser Vorteil von Kontrollvariablen kommt insbesondere bei komplexen Versuchsumgebungen und Aufgaben sowie kleinen Stichproben zum Tragen (Aguinis, Beaty, Boik, & Pierce, 2005; Jo, 2002). Daher sollten in zukünftigen Untersuchungen zur Wirkungsüberprüfung einer Variablen weiterhin potentielle Störeinflüsse theoriegeleitet kontrolliert oder ausgeschaltet werden.

### ***4. Optimale Stichprobengrößen***

In den Untersuchungen von SAM 1.5 (s. Abschnitt 3.2.1 und 3.2.2) sowie der Leistungs-determinante „Informationsverfügbarkeit“ (s. Abschnitt 5.2.3) konnten für die Aufklärung einiger zu überprüfender Aspekte nur unzureichende statistische Validitäten erreicht werden. Dies war durch den Einsatz der speziellen VPn (Extremausprägung hinsichtlich unspezifischer Strategie bzw. professionelle Entwickler von Automationen) bedingt, von denen nur wenige gewonnen werden konnten. Dadurch wiesen die statistischen Analysen teilweise, trotz gefundener mittlerer Effekte zu geringe Teststärken für eine statistische Absicherung auf. Mittlere Effekte werden von der Autorin jedoch als praktisch bedeutsam eingestuft. Deshalb sollte in nachfolgenden Untersuchungen verstärkt auf eine optimale Stichprobengröße geachtet werden, um eine statistische Absicherung gefundener Effekte gewährleisten zu können (s. Bortz & Döring, 2006).

### ***5. Weiterentwicklung des „Bewertungsinstruments der Güte von Automatikkonzepten“***

Ebenso wie das ATEO Lab System (s. Abschnitt 3.3) stellt das „Bewertungsinstrument der Güte von Automatikkonzepten“ (s. Abschnitt 5.2.2) ein nützliches Werkzeug für weitere Forschungstätigkeiten in diesem sowie anderen Bereichen dar. Durch dieses konnten Unterschiede zwischen den Automatikkonzepten aus der Untersuchung der Leistungs-determinante „Informationsverfügbarkeit“ identifiziert werden. Für weitere Einsätze des Instruments sollten die Bewertungskriterien anwendungsökonomischer operationalisiert und durch Beispiele veranschaulicht werden. Zugleich sollte das Instrument um die Bewertung von Implementationen von Automationen zur Prozessüberwachung und -führung von komplexen, dynamischen Systemen erweitert werden. Dabei sollte weiterhin ein modular aufgebautes Instrument angestrebt werden, bei welchem, je nach Anwendungsanforderung, einzelne Module für die Beurteilung ausgewählt werden können. Ferner sollten die Bewertungskriterien in einem Workshop mit Experten hinsichtlich deren Bedeutsamkeit beurteilt und die Gütekri-

terien des Instruments bestimmt werden. Um den Aufwand bei der Anwendung trotz der Erweiterung zu minimieren, wäre eine Umwandlung der XLS-Tabelle in eine Software von Vorteil, in welche die ei

nzlenen Aspekte eines Konzepts nach bestimmten Kriterien zur automatischen Auswertung eingegeben werden. Dadurch kann eine hohe Beurteilerkonkordanz erreicht werden.

## **6. Experimenteller Vergleich Entwickler und Operateur**

Im Anschluss an diese Dissertation war der nächste Schritt der Vergleich zwischen Entwicklern und Operateuren hinsichtlich deren Leistungsfähigkeit bei der Überwachung und Führung von SAM unter den optimalen Ausprägungen der identifizierten jeweiligen Leistungsdeterminanten. Dieser Vergleich wurde von Niestroj (2014) in der dritten ATEO-Projektphase durchgeführt. Dafür wurde das Automatikkonzept mit der höchsten Güte und besten Umsetzbarkeit aus der Untersuchung der Leistungsdeterminante „Informationsverfügbarkeit“ (s. Abschnitt 5.2.3) vollständig in die AA-S implementiert (s. Abschnitt 3.3.1). Das Automatikkonzept des Teams 37 der Bedingung 3 weist die höchste Güte auf (s. Anhang G.6.6) und sollte daher ursprünglich für diesen experimentellen Vergleich herangezogen werden. Allerdings zeigte sich im Laufe seiner Implementierung, dass einzelne Funktionen nicht durch die Programmierumgebung Squeak umsetzbar waren. Daher wurde von Niestroj (2014) das Konzept 36 ausgewählt, da dieses die zweithöchste Güte aufweist und dieses einfacher umsetzbar ist. Dieser konnte durch den Vergleich nachweisen, dass die Automaten des ausgewählten Konzepts zu einer signifikant höheren Genauigkeit des SAM-Prozesses führen, als es ein Operateur umzusetzen vermag. Ferner wurde in dieser ATEO-Projektphase ein Schritt weitergegangen und die Thematik „kooperative Automatisierung“ (s. Hakuli, et al., 2009) experimentell beleuchtet (s. Niestroj, 2014; von Bernstorff, 2014). Dabei zeigte sich, dass ein Operateur, unterstützt durch kooperative Automationen, ebenfalls eine höhere Genauigkeit des SAM-Prozesses erreicht, als es einem nicht unterstützten Operateur möglich ist (s. Niestroj, 2014). Zugleich konnten Empfehlungen für eine flexible Funktionsallokation in realen komplexen, dynamischen, technischen Systemen gewonnen werden (s. von Bernstorff, 2014). Abschließend sei erwähnt, dass die Autorin einen Erfahrungs- und Informationsaustausch für Forschungstätigkeiten anbietet, die an diese Arbeit oder das Projekt ATEO anknüpfen.“

## Literaturverzeichnis

- DIN EN 894: Ergonomische Anforderungen an die Gestaltung von Anzeigen und Stellteilen. Deutsches Institut für Normung e. V. Berlin: Beuth (1997).
- ISO/IEC 9126: Software Engineering - Product Quality. (1991).
- DIN EN ISO 9241-110: Ergonomische Anforderungen der Mensch-System-Interaktion - Teil 110: Grundsätze der Dialoggestaltung. Deutsches Institut für Normung e. V. Berlin: Beuth. (2006).
- DIN ISO/IEC 12199: Software-Erzeugnisse - Qualitätsanforderungen und Prüfbestimmungen. Deutsche Fassung IST/IEC 12119. Berlin: Beuth-Verlag (1995).
- DIN EN ISO 15005: Ergonomische Aspekte von Fahrerinformations- und -assistenzsystemen. Grundsätze und Prüfverfahren des Dialogmanagements. Deutsches Institut für Normung e. V. Berlin: Beuth. (2003).
- DIN IEC 60050-351:2009-06: Internationales Elektrotechnisches Wörterbuch - Teil 351: Leittechnik. (2009).
- Acuña, S. T., Gómez, M., & Juristo, N. (2009). How do personality, team processes and task characteristics relate to job satisfaction and software quality? *Information and Software Technology*, 51(3), 627-639.
- Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: a 30-year review. *Journal of Applied Psychology*, 90(1), 94-107.
- Ahmed, S., Wallace, K. M., & Blessing, L. T. M. (2003). Understanding the Differences between How Novice and Experienced Designers Approach Design Tasks. *Research in Engineering Design*, 14(1), 1-11.
- Annett, J., Duncan, K. D., Stammers, R. B., & Gray, M. J. (1971). *Task Analysis*. Department of Employment Training Information Paper No. 6. London: Her Majesty's Stationary Office (HMSO).
- Anderson, J. R. (2001). *Kognitive Psychologie*. Heidelberg: Spektrum.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64(6) 359-372.



- Auer, P., & von der Weth, R. (1994). Wie klären Konstrukteure ein Problem? - Die Entwicklung von Erfahrung und das Vorgehen bei der Analyse von Konstruktionsaufgaben. *46*(5), 175-180.
- Baber, C., & Wankling, J. (1992). An experimental comparison of text and symbols for in-car reconfigurable displays. *Applied Ergonomics*, *23*(4), 255-262.
- Badke-Schaub, P. (2005). Design performance: how can we meet human limitations with human resources? In A. Samuel & W. Lewis (Eds.), *Engineering Design and the Global Economy* (pp. 188-189). Australia: Institution of Engineers.
- Badke-Schaub, P., & Frankenberger, E. (2004). *Management kritischer Situationen: Produktentwicklung erfolgreich gestalten*. Berlin: Springer Verlag.
- Badke-Schaub, P., Hofinger, G., & Lauche, K. (2008). *Human Factors. Psychologie sicheren Handelns in Risikobranchen*. Heidelberg: Springer.
- Badke-Schaub, P., & Stempfle, J. (2003). Analysis of solution finding processes in design teams. In U. Lindemann (Ed.), *Human Behavior in Design – Individuals, Teams, Tools* (pp. 121-131). Berlin: Springer Verlag.
- Bainbridge, L. (1983). Ironies of Automation. *Automatica*, *19*(6), 775-779.
- Bainbridge, L. (1987). Ironies of Automation. In J. Rasmussen, K. Duncan & J. Leplat (Eds.), *New technology and human error* (pp. 271-283). Chichester: Wiley.
- Bainbridge, L. (1998). Complex Cognition - Papers by Lisanne Bainbridge. Retrieved 24. February, 2015, from <http://www.bainbrdg.demon.co.uk/index.html>
- Balzert, H. (2000). *Lehrbuch der Software-Technik. Band 1: Software-Entwicklung* (2nd ed.). Heidelberg: Spektrum Akademischer Verlag.
- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: on the automaticity of higher mental processes. *Psychological Bulletin*, *126*(6), 925-945.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five Personality Dimension and Job Performance: A Meta-Analysis. *Personnel Psychology*, *44*(1), 1-26.
- Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology*, *83*(3), 377-391.

- Baya, V., & Leifer, L. J. (1996). Understanding Information Management in Conceptual Design. In N. Cross, H. Christiaans & K. Dorst (Eds.), *Analyzing Design Activity* (pp. 151-168). Chichester: John Wiley & Sons.
- Becker-Beck, U. (1994). Strukturanalyse des Interaktionsverhaltens in Diskussionsgruppen. *Gruppendynamik*, 25(1), 95-106.
- Bell, S. T. (2004). Setting the stage for effective teams: A meta-analysis of team design variable and team effectiveness. Unpublished Dissertationsschrift. Texas A&M University.
- Bell, S. T. (2007). Deep-Level Composition Variables as Predictors of Team Performance: A Meta-Analysis. *Journal of Applied Psychology*, 92(3), 595-615.
- Belz, S. M., Robinson, G. S., & Casali, J. G. (1999). A New Class of Auditory Warning Signals for Complex Systems: Auditory Icons. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 41(4), 608-618.
- BildschArbV: Verordnung über Sicherheit und Gesundheitsschutz bei der Arbeit an Bildschirmgeräten, Umsetzung der EWGRL 270/90 vom 4. Dezember 1996, BGB I 1996, 1843 (1996).
- Billings, C. E. (1997). *Aviation Automation: The Search for a Human-Centered Approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- BIT. (2015). V-Modell XT. Retrieved 23. December, 2015, from [http://www.cio.bund.de/Web/DE/Architekturen-und-Standards/V-Modell-XT/vmodell\\_xt\\_node.html](http://www.cio.bund.de/Web/DE/Architekturen-und-Standards/V-Modell-XT/vmodell_xt_node.html)
- Blau, G. (1993). Operationalizing Direction and Level of Effort and Testing Their Relationships to Individual Job Performance. *Organizational Behavior and Human Processes*, 55(1), 152-170.
- Boehm, B. W. (1988). A Spiral Model of Software Development and Enhancement. *IEEE Computer*, 21(5), 61-72.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler* (5th ed.). Berlin: Springer-Verlag.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6th ed.). Berlin: Springer-Verlag.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4th ed.). Berlin: Springer Verlag.

- Bortz, J., & Lienert, G. A. (1998). *Kurzgefaßte Statistik für die klinische Forschung - Ein praktischer Leitfaden für die Analyse kleiner Stichproben*. Berlin: Springer-Verlag.
- Bothe, K., & Hildebrandt, M. (2008). *ATEO - Verhaltensspezifikation mit Erweiterungswünschen* (Forschungsbericht): Humboldt Universität zu Berlin, Institut für Informatik.
- Bothe, K., Hildebrandt, M., & Niestroj, N. (2009). ATEO-System Komponente: SAMS. Unpublished Verhaltensspezifikation. Humboldt-Universität zu Berlin, Institut für Informatik.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81(3), 211-241.
- Brehmer, B. (2005). Micro-worlds and the circular relation between people and their environment. *Theoretical Issues in Ergonomics Science*, 6(1), 73-93.
- Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior* 9(2-3), 171-184.
- Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In S. Ashlund, A. Henderson, E. Hollnagel, K. Mullet & T. White (Eds.), *Proceedings of the INTERCHI '93 conference on Human factors in computing systems* (pp. 222-227). Amsterdam: ACM Press.
- Brodbeck, F. C. (2001). Communication and performance in software development projects. *European Journal of Work and Organizational Psychology*, 10(1), 73-94.
- Brown, S. L., & Eisenhardt, K. M. (1995). Product Development: Past Research, Present Findings, and Future Directions. *The Academy of Management Review*, 20(2), 343-378.
- Bühl, A., & Zöfel, P. (2005). *SPSS 12: Einführung in die moderne Datenanalyse unter Windows* (9th ed.). München: Pearson Studium.
- Burandt, M. (2007). ATEO - begleitende Arbeit zur Projektdurchführung. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

- Burkhard, H. D., Jahn, L., Kain, S., Meyer, C., Muetterlein, J., Nachtwei, J., et al. (2011). Artificial Subjects in Psychological Experiments based on the "Socially Augmented Microworld (SAM)". In M. Szczuka, L. Czaja, A. Skowron & M. Kacprzak (Eds.), *Proceedings of the international workshop CS&P '2011*, 28-30 September 2011 (pp. 54-65). Poland: Pultusk.
- Burmester, M., Graf, R., Hellbrück, J., & Meroth, A. (2008). Usability - Der Mensch im Fahrzeug. In A. Meroth & B. Tolg (Eds.), *Infotainmentsysteme im Kraftfahrzeug: Grundlagen, Komponenten, Systeme und Anwendungen* (pp. 321-355). Wiesbaden: Vieweg Verlag.
- Burmester, M., Komischke, T., & Wüst, L. (2000). Innovative User Interfaces in Automation Engineering by Application of Usability Engineering Methods Shown by the Example of a Three-Dimensional Plant Representation. *INTERNATIONAL JOURNAL OF HUMAN-COMPUTER INTERACTION*, 12(3-4), 359-373.
- Burns, C. M., & Vicente, K. M. (1995). A framework for describing and understanding interdisciplinary interactions in design. In G. M. Olson & S. Schuon (Eds.), *Proceedings of the 1st conference on designing interactive systems: processes, practices, methods, and techniques* (pp. 97-103). 23-25 August 1995, Ann Arbor, Michigan. New York: ACM Press.
- Byrne, Z. S., Mueller-Hanson, R. A., Cardador, J. M., Thornton III, G. C. I., Schuler, H., Frintrup, A., et al. (2004). Measuring achievement motivation: tests of equivalency for English, German, and Israeli versions of the Achievement Motivation Inventory. *Personality and Individual Differences*, 37(1), 203-217.
- Campbell, J. L., Richard, C. M., Brown, J. L., & McCallum, M. (2007). *Crash warning system interfaces: Human factors insights and lessons learned* (No. DOT HS 810697). Seattle: Battelle Center for Human Performance and Safety.
- Campbell, J. L., Richmann, J. B., Carney, C., & Lee, J. D. (2004). *In-vehicle display icons and other information elements: Preliminary assessment of visual symbols. Volume 1: Guideline* (No. FHWA-RD-03-065): U.S. Department of Transportation.
- Christiaans, H. H. C. M., & Dorst, K. H. (1992). Cognitive Models in Industrial Design Engineering: A Protocol Study. In D. L. Taylor & D. A. Stauffer (Eds.), *Design theory and methodology DTM'92* (pp. 131-141). New York: American Society of Mechanical Engineers.

- Clamann, M. P., Wright, M. C., & Kaber, D. B. (2002). Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing. *In Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, 342-346.
- Costa Jr., P. T., & McCrae, R. R. (1997). Longitudinal stability of adult personality. In R. Hogan, J. Johnson & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 269-290). San Diego, CA: Academic Press.
- Cotter, S., Hopkin, J., Stevens, A., Burrows, A., Kompfner, P., & Flanment, M. (2006). The institutional context for Advanced Driver Assistance Systems: A code of practice for developement. *In Proceedings of the 13th ITS World Congress*. London 8-12 October 2006. Crowthorne Berks, UK: TRL Library.
- Cropley, A. (2006). In Praise of Convergent Thinking. *Creativity Research Journal*, 18(3), 391-404.
- Cross, N., Christiaans, H., & Dorst, K. (1996). *Analysing Design Activity*. Chichester: John Wiley & Sons.
- Cummings, M. L., & Thornburg, K. M. (2011). Paying Attention to the Man behind the Curtain. *IEEE Pervasive Computing*, 10(1), 58-62.
- Diaper, D., & Stanton, N. (2004). *The handbook of task analysis for human-computer interaction*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Diefenbach, S., Chien, W.-C., Lenz, E., & Hassenzahl, M. (2013). Prototypen auf dem Prüfstand. Bedeutsamkeit der Repräsentationsform im Rahmen von Konzeptevaluation. *icom. Zeitschrift für interaktive und kooperative Medien*, 12(1), 53-63.
- DiFonzo, N., Hantula, D. A., & Bordia, P. (1998). Microworlds for experimental research: Having your (control and collection) cake, and realism too. *Behavior Research Methods, Instruments, & Computers*, 30(2), 278-286.
- Dirkin, G. R. (1983). Cognitive tunneling: use of visual information under stress. *Perceptual and Motor Skills*, 56(1), 191-198.
- Doane, S. M., Sohn, Y. W., & Jodlowski, M. (2004). Pilot Ability to Anticipate the Consequences of Flight Actions as a Function of Expertise. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 46(1), 92-103.

- Dodt, K. (2012). Anwendung des IKD (Instrument zur Kodierung von Diskussionen) zur Ermittlung des Einflusses von sozialen Prozessen auf den Konzeptionsprozess und das Ergebnis. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Dörner, D. (1987). *Problemlösen als Informationsverarbeitung* (3rd ed.). Stuttgart: Kohlhammer.
- Dörner, D. (1989). *Die Logik des Misslingens. Strategisches Denken in komplexen Situationen*. Reinbek bei Hamburg: Rowohlt Verlag.
- Dörner, D. (1995). Problemlösen und Gedächtnis. In D. Dörner & E. v. d. Meer (Eds.), *Das Gedächtnis. Probleme – Trends – Perspektiven* (pp. 295-320). Göttingen: Hogrefe.
- Dorst, K. (2003). The Problem of Design Problems. In N. G. Cross and E. Edmonds (eds.), *Expertise in Design: design thinking research symposium 6* (135-147). Sydney, Australia: Creativity and Cognition Studios Press.
- Dullat, S. (n. D.). Optimierungsmöglichkeiten der Performance einer softwarebasierten Versuchsumgebung. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Dylla, N. (1991). *Denk- und Handlungsabläufe beim Konstruieren. Konstruktionstechnik*. (Band 5). München: Hanser.
- Edworthy, J., & Adams, A. (1996). *Warning Design: A Research Prospective*. London: Taylor & Francis.
- Ehrlenspiel, K. (1993). Denkfehler bei der Maschinenkonstruktion: Beispiele, Gründe und Hintergründe. In S. Strohschneider & R. von der Weth (Eds.), *Ja, mach nur einen Plan* (pp. 196-207). Bern: Huber.
- Ehrlenspiel, K. (2003). *Integrierte Produktentwicklung: Methoden für Prozeßorganisation, Produkterstellung und Konstruktion* (2nd ed.). München: Hanser.
- Ehrlenspiel, K., Bernard, R., & Günther, J. (1996). Unterstützung des Konstruktionsprozesses durch Modelle. In W. Hacker & P. Sachse (Eds.), *Bild und Begriff III. Zur Rolle von Anschauung und Abstraktion im konstruktiven Entwurfsprozess. Bericht über das Werkstattgespräch in Diesbar-Seußlitz vom 20.-22.10.1995* (pp. 45-59). Dresden: TU-Eigenverlag.

- Ehrlenspiel, K., & Günther, J. (1995). Wie wird Konstruieren erfolgreich? In Verein Deutscher Ingenieure (Ed.), *Effizienter entwickeln und konstruieren. VDI Berichte 1169* (pp. 45-70). Düsseldorf: VDI.
- Ehrlenspiel, K., Kiewert, A., & Lindemann, U. (2005). *Kostengünstig Entwickeln und Konstruieren: Kostenmanagement bei der integrierten Produktentwicklung* (5th ed.). Berlin: Springer-Verlag.
- Eilers, K., Nachreiner, F., & Hänecke, K. (1986). Entwicklung und Überprüfung einer Skala zur Erfassung subjektiv erlebter Anstrengung. *Zeitschrift für Arbeitswissenschaft*, 40(4), 215-224.
- Eisentraut, R., & Günther, J. (1997). Individual styles of problem solving and their relation to representations in the design process. *Design Studies*, 18(4), 369-383.
- Ellis, P. D. (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge, New York: Cambridge University Press.
- Emmanuelides, P. A. (1993). Towards an integrative framework of performance in product development projects. *Journal of Engineering and Technology Management*, 10(4), 363-392.
- Endsley, M. R. (1987). The application of human factors to the development of expert systems for advanced cockpits. *Proceedings of the Human Factors Society - 31st Annual Meeting (Santa Monica, CA: Human Factors Society)*, 1388-1392.
- Endsley, M. R. (1995a). Measurement of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 65-84.
- Endsley, M. R. (1995b). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 32-64.
- Endsley, M. R. (1997). *Situation Awareness, Automation & Free Flight*. Paper presented at the Presentation to the FAA/Eurocontrol Air Traffic Management R&D Seminar. June, Saclay, France.
- Endsley, M. R., Bolté, B., & Jones, D. G. (2003). *Designing for Situation Awareness: An Approach to User-Centered Design*. London: CRC Press Taylor & Francis Group.

- Endsley, M. R., Hoffman, R., Kaber, D. B., & Roth, E. (2007). Cognitive Engineering and Decision Making: An Overview and Future Course. *Journal of Cognitive Engineering and Decision Making*, 1(1), 1-21.
- Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42(3), 462-492.
- Endsley, M. R., & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(2), 381-394.
- Eris, O. (2003). Manifestation of Divergent-Convergent Thinking in Question Asking and Decision Making Processes of Design Teams: A Performance Dimension. In U. Lindemann (Ed.), *Human Behavior in Design: Individuals, Teams, Tools* (pp. 142-153). Berlin: Springer Verlag.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- Fitts, P. M. (1951). *Human engineering for an effective air-navigation and traffic control system* (Report for the Air Navigation Development Board). Columbus: Ohio State University Research Foundation.
- Flemisch, F. O., Kelsch, J., Löper, L., Schieben, A., & Schindler, J. (2008). Automation spectrum, inner / outer compatibility and other potentially useful human factors concepts for assistance and automation. In D. de Waard, F.O. Flemisch, H. B. Lorenz, H. Oberheid & K. A. Brookhuis (Eds.), *Human Factors for assistance and automation* (pp. 257-272). Maastricht, the Netherlands: Shaker Publishing.
- Frank, P., & Reichelt, W. (2001). Fahrerassistenzsysteme im Entwicklungsprozess. In T. Jürgensohn & K.-P. Timpe (Eds.), *Kraftfahrzeugführung* (pp. 71-78). Berlin: Springer Verlag.
- Frankenberger, E. (1997). *Arbeitsteilige Produktentwicklung: Empirische Untersuchung und Empfehlungen zur Gruppenarbeit in der Konstruktion*. Dissertation Technische Universität Darmstadt. Düsseldorf: VDI-Verlag; VDI-Fortschrittsberichte, Nr. 291.
- Fricke, G. (1993). *Konstruieren als flexibler Problemlöseprozess: Empirische Untersuchung über erfolgreiche Strategien und methodische Vorgehensweisen beim Konstruieren*. Dissertation Technische Universität Darmstadt. Düsseldorf: VDI-Verlag; VDI-Fortschrittsberichte, Nr. 227



- Fricke, G. (1994). Erfolgreiches individuelles Vorgehen beim Konstruieren: Ergebnisse einer empirischen Untersuchung. *Konstruktion*, 46, 181-189.
- Fricke, N. (2009). Gestaltung zeit- und sicherheitskritischer Warnungen im Fahrzeug. Unpublished Dissertation. Technische Universität Berlin, Fakultät V - Verkehrs- und Maschinensysteme.
- Fricke, N., De Filippis, M., & Thüring, M. (2008). Wie beurteilt man sicherheitskritische Fahrerassistenzsysteme? – Darstellung einer Fahrsimulatorstudie. *Zeitschrift für Arbeitswissenschaft*, 62(1), 25-35.
- Frieling, E., & Sonntag, K. (1999). *Lehrbuch der Arbeitspsychologie* (2nd ed.). Bern: Hans Huber.
- Fuhrmann, E. (2010). Entwicklung eines GUI für die Konfiguration der Software-Komponente zur Systemprozessüberwachung und -kontrolle in einer psychologischen Versuchsumgebung. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Fuld, R. B. (1993). The Fiction of Function Allocation. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 1(1), 20-24.
- Galster, S. M., Bolia, R. S., & Parasuraman, R. (2002). Effects of Information Automation and Decision-Aiding Cuing on Action Implementation in a Visual Search Task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting September 2002*, 46(3), 438-442.
- Gebert, D. (2004). *Innovation durch Teamarbeit: Eine kritische Bestandsaufnahme*. Stuttgart: Kohlhammer.
- Gérard, N., Huber, S., Nachtwei, J., Satriadarma, B., & Schubert, U. (2011). A Framework for Designers to Support Prospective Design of Human Computer Interaction. *International Journal on Human-Computer Interaction*, 2(7), 17-38.
- Glowinski, D., Mancini, M., Cowie, R., Camurri, A., Chiorri, C., & Doherty, C. (2013). The movements made by performers in a skilled quartet: a distinctive pattern, and the function that it serves. *Frontiers in Psychology*, 4, 1-9.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of Personality and Social Psychology* (Vol. 2, pp. 141-165). Beverly Hills: CA: Sage.

- Gonzalez, C., Vanyukov, P., & Martin, M. K. (2005). The use of microworlds to study dynamic decision making. *Computers in Human Behavior*, 21, 273-286.
- Goodman, J., Clarke, S., Langdon, P., & Clarkson, P. J. (2007). Designers' Perceptions of Methods of Involving and Understanding Users. In C. Stephanidis (Ed.), *Universal Access in HCI* (Band 4554, pp. 127-136). Heidelberg: Springer-Verlag.
- Görner, R. (1994). Zur psychologischen Analyse von Konstrukteur- und Entwurfstätigkeiten. In B. Bergmann & P. Richter (Eds.), *Die Handlungsregulationstheorie. Von der Praxis einer Theorie* (pp. 233-241). Göttingen: Hogrefe.
- Graham, R. (1999). Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics*, 42(9), 1233-1248.
- Gross, B. (2006). Mikrowelten in Squeak. In A. Heinecke & H. Paul (Eds.), *Mensch und Computer 2006. Workshop-Beiträge zur sechsten fachübergreifenden Konferenz* (pp. 59-64). München: Oldenbourg Wissenschaftsverlag.
- Gross, B. (2008). *Gestaltung von Schwierigkeitsgraden einer Versuchsumgebung: Der Ansatz der belebten Mikrowelt SAM*. Saarbrücken: VDM Verlag Dr. Müller.
- Gross, B., & Nachtwei, J. (2006). Assistenzsysteme effizient entwickeln und nutzen - Die Mikrowelt als Methode zur Wissensakquisition für Entwickler und Operateure. In M. Grandt & A. Bauch (Eds.), *Cognitive Systems Engineering in der Fahrzeug- und Prozessführung* (pp. 75-88). Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt e. V. (DGLR-Bericht 2006-02/07).
- Gross, B., & Nachtwei, J. (2007). How to develop and use assistance systems efficiently - Using the microworld as a method to acquire knowledge for developers and operators. In D. de Waard, B. Hockey, P. Nickel & K. Brookhuis (Eds.), *Human Factors Issues in Complex System Performance* (pp. 345-350). Maastricht, the Netherlands: Shaker Publishing.
- Grote, G., Weik, S., Wäfler, T., Zölch, M., & Ryser, C. (1999). KOMPASS: Komplementäre Analyse und Gestaltung von Produktionsaufgaben in soziotechnischen Systemen: Eine arbeitspsychologische Methode zur Optimierung der Mensch-Maschine-Funktionsteilung. In H. Dunckel (Ed.), *Handbuch psychologischer Arbeitsanalyseverfahren. Mensch, Technik, Organisation* (Bd 14, pp. 255-284). Zürich: vdf Hochschulverlag an der ETH Zürich.

- Grubert, J. (n. D.). ungewichtete deutsche Kurzversion des NASA Task Load. Retrieved 11. July, 2015, from <http://jensgrubert.bplaced.net/nasa-tlx-short/TLX-German-short.html>
- Gründl, M. (2005). Fehler und Fehlverhalten als Ursache von Verkehrsunfällen und Konsequenzen für das Unfallvermeidungspotenzial und die Gestaltung von Fahrerassistenzsystemen. Unpublished Dissertation. Universität Regensburg.
- Günther, J. (1998). Individuelle Einflüsse auf den Konstruktionsprozess: Eine empirische Untersuchung unter besonderer Berücksichtigung von Konstrukteuren aus der Praxis. Aachen: Shaker.
- Günther, J., & Ehrlenspiel, K. (1998). How do designers from practise design? In H. Birkhofer, P. Badke-Schaub & E. Frankenberger (Eds.), *Designers: The Key to Successful Product Development* (pp. 85-97). London: Springer Verlag.
- Guzdial, M. (2001). *Squeak: Object-oriented design with multimedia applications*. Englewood Cliffs: NJ Prentice Hall.
- Haas, E., & Edworthy, J. (2006). An introduction to auditory warnings and alarms. In M. S. Wogalter (Ed.), *Handbook of Warnings* (pp. 189-198). Mahwah: Lawrence Erlbaum Associates.
- Hacker, W. (1989). On the utility of procedural rules: Conditions of the use of rules in the production of operation sequences. *Ergonomics*, 32(7), 717-732.
- Hacker, W. (1994). Zur Struktur von Handlungswissen bei Entwurfstätigkeiten. In G. Pahl (Ed.), *Psychologische und pädagogische Fragen beim methodischen Konstruieren. Ergebnisse des Ladenburger Diskurses vom Mai 1992 bis Oktober 1993* (pp. 161-175). Köln: TÜV Rheinland.
- Hacker, W. (1996). Entwickeln und Konstruieren als Denktätigkeit: zu einer Arbeitswissenschaft geistiger Erwerbstätigkeit. *Zeitschrift für Arbeitswissenschaft*, 50(2), 111-116.
- Hacker, W. (1998). *Allgemeine Arbeitspsychologie: Psychische Regulation von Arbeitstätigkeiten*. Bern: Huber.
- Hacker, W., & Sachse, P. (2006). Entwurfstätigkeiten und ihre psychologischen Unterstützungsmöglichkeiten. In B. Zimolong & U. Konradt (Eds.), *Ingenieurpsychologie* (pp. 671-707). Göttingen: Hogrefe.

- Hacker, W., & Wetzstein, A. (2004). Verbalisierende Reflexion und Lösungsgüte beim Entwurfsdenken. *Zeitschrift für Psychologie*, 212(3), 152-166.
- Hacker, W., Wetzstein, A., & Römer, A. (2002). Gibt es Vorgehensmerkmale erfolgreichen Entwerfens von Produkten? *Zeitschrift für Arbeitswissenschaft*, 5, 305-317.
- Hackman, J. R. (1987). The design of work teams. In J. W. Lorsch (Ed.), *Handbook of organizational behaviour* (pp. 315-342). Prentice-Hall, NJ: Englewood Cliffs.
- Hakuli, S., Bruder, R., Flemisch, F., Löper, C., Rausch, H., Schreiber, M., et al. (2009). Kooperative Automation. In H. Winner, S. Hakuli & G. Wolf (Eds.), *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (pp. 647-656). Wiesbaden: Vieweg + Teubner.
- Hampel, T. (2012). Konzeption und Implementation eines Analysetools zur Auswertung von Logfiles einer komplexen, dynamischen Versuchsumgebung. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Harrison, D. A., Mohammed, S., McGrath, J. E., Florey, A. T., & Vanderstoep, S. W. (2003). Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality. *Personnel Psychology*, 56(3), 633-669.
- Hart, S. G., & Staveland, L. E. (1988). Development of a multidimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (Vol. 52, pp. 139-183). Amsterdam: Elsevier.
- Hasselmann, M. (2010). ATEO: Der Operatorsarbeitsplatz in mobiler Umgebung. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Hasselmann, M. (n. D.). Erweiterung einer Softwarekomponente für Automaten zur Systemprozessüberwachung und -führung als Bestandteil einer psychologischen Versuchsumgebung. Diplomarbeit in Vorbereitung. Humboldt-Universität zu Berlin, Institut für Informatik.
- Hauß, Y., & Timpe, K.-P. (2000). Automatisierung und Unterstützung im Mensch-Maschine-System. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Eds.), *Mensch-Maschine-Systemtechnik: Konzepte, Modellierung, Gestaltung, Evaluation* (pp. 41-62). Düsseldorf: Symposion Publishing.
- Hays, R. T., & Singer, M. J. (1989). *Simulation Fidelity in Training System Design: Bridging the Gap Between Reality and Training*. New York: Springer.

- Hedicke, V. (2000). Multimodalität in Mensch-Maschine-Schnittstellen. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Eds.), *Mensch-Maschine-Systemtechnik: Konzepte, Modellierung, Gestaltung, Evaluation* (pp. 203-232). Düsseldorf: Symposion Publishing.
- Henning, M. J., Kienast, H., Lindner, P., Weigel, H., Krems, J. F., & Spanner-Ulmer, B., et al. (2008). Multifunktionale Mensch-Maschine-Schnittstelle. In M. Maurer & C. Stiller (Eds.), *5. Workshop Fahrerassistenzsysteme FAS 2008* (pp. 151-158). Karlsruhe: fmrt e.V.
- Herczeg, M. (1994). *Software-Ergonomie: Grundlagen der Mensch-Computer-Kommunikation*. Bonn: Addison-Wesley Verlag.
- Herczeg, M. (2009). *Software-Ergonomie: Theorien, Modelle und Kriterien für gebrauchstaugliche interaktive Computersysteme* (3rd ed.). München: Oldenbourg Wissenschaftsverlag GmbH.
- Herzberg, P. Y., & Brähler, E. (2006). Assessing the Big-Five Personality Domains via Short Forms: A Cautionary Note and a Proposal. *European Journal of Psychological Assessment*, 22(3), 139-148.
- Hilburn, B., Jorna, P., Byrne, E. A., & Parasuraman, R. (1997). The effect of adaptive air traffic control (ATC) decision aiding on controller mental workload. In M. Mouloua & J. M. Koonce (Eds.), *Human-automation interaction: Research and practice* (pp. 84-91). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hilburn, B., Molloy, R., Wong, D., & Parasuraman, R. (1993). Operator versus computer control of adaptive automation. In J. G. Morrison (Ed.), *The Adaptive Function allocation for Intelligent Cockpits (AFAIC) program: Interim research and guidelines for the application of adaptive automation* (Progress Report. No. NAWCADWAR-93031-60., pp. 31-36). Warminster, PA: Air Vehicle and Crew Systems Technology Department.
- Hildebrandt, M. (2008). Analyse und Evaluierung der Architektur des ATEO-Systems. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Hildebrandt, M. (2009). Reengineering der Architektur der SAM-Komponente des ATEO-Systems und deren Prototypische Implementation in Java. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

- Hildebrandt, M., Kain, S., Kesselring, K., Nachtwei, J., Niestroj, N., & Schwarz, H. (2010). Die Hierarchische Aufgabenanalyse im Babel interdisziplinärer Softwareentwicklungsprojekte. *Zeitschrift für Arbeitswissenschaft*, 04, 374-380.
- Ho, C., & Spence, C. (2009). Using peripersonal warning signals to orient a driver's gaze. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 51(4), 539-556.
- Hoc, J.-M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *International Journal of Human Computer Studies*, 54(4), 509-540.
- Högl, M. (1998). *Teamarbeit in innovativen Projekten. Einflussgrößen und Wirkungen*. Wiesbaden: Deutscher Universitäts-Verlag.
- Hollister, W. M. (1986). *Improved Guidance and Control Automation at the Man-Machine Interface*. Neuilly sur Seine, France: Advisory Group for Aerospace Research & Development.
- Holtzblatt, K., & Hugh, R. B. (2014). Contextual Design. In M. Soegaard & R. F. Dam (Eds.), *The Encyclopedia of Human-Computer Interaction* (2nd ed., Chapter 8). Retrieved 01. January, 2016, from <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed>
- Horrey, W., Wickens, C., & Consalus, K. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Journal of Experimental Psychology: Applied*, 12(2), 67-78.
- Huber, S., Kain, S., & Nachtwei, J. (2008). Effekte sicherer nachweisen: Persönlichkeitsmerkmale als Kontrollvariablen in der Human Factors Forschung. In M. Grandt & A. Bauch (Eds.), *Beiträge der Ergonomie zur Mensch-System-Integration. Deutsche Gesellschaft für Luft- und Raumfahrt e.V. (DGLR-Bericht 2008, pp. 143-159)*. Bonn.
- Ingalls, D., Kaehler, T., Maloney, J., Wallace, S., & Kay, A. (1997). Back to the future: the story of Squeak, a practical Smalltalk written in itself. *OOPSLA '97: Proceedings of the 12th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications*. (pp. 318-326). New York: ACM.
- Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods*, 7(2), 178-193.

- Johannsen, G. (1993). *Mensch-Maschine-Systeme*. Berlin: Springer Verlag.
- Johansson, B., Persson, M., Granlund, R., & Mattsson, P. (2003). C3Fire in command and control research. *Cognition, Technology & Work*, 5(3), 191-196.
- Kaber, D. B., & Endsley, M. R. (1997). Out-of-the-Loop Performance Problems and the Use of Intermediate Levels of Automation for Improved Control System Functioning and Safety. *Process Safety Progress*, 16(3), 126-131.
- Kaber, D. B., & Endsley, M. R. (1998). Team Situation Awareness for Process Control Safety and Performance. *Process Safety Progress*, 17(1), 43-48.
- Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science*, 5(2), 113-153.
- Kaber, D. B., Onal, E., & Endsley, M. R. (2000). Design of Automation for Telerobots and the Effect on Performance, Operator Situation Awareness, and Subjective Workload. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 10(4), 409-430.
- Kaber, D. B., & Riley, J. M. (1999). Adaptive Automation of a Dynamic Control Task Based on Secondary Task Workload Measurement. *International Journal of Cognitive Ergonomics*, 3(3), 169-187.
- Kaber, D. B., Riley, J. M., Tan, K.-W., & Endsley, M. R. (2001). On the Design of Adaptive Automation for Complex Systems. *International Journal of Cognitive Ergonomics*, 5(1), 37-57.
- Kaber, D. B., Stoll, N., Thurow, K., Green, R. S., Kim, S.-H., & Mosaly, P. (2009). Human-Automation Interaction Strategies and Models for Life Science Applications. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 19(6), 601-621.
- Kaber, D. B., Wright, M. C., Prinzel III, L. J., & Clamann, M. P. (2005). Adaptive Automation of Human-Machine System Information-Processing Functions. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(4), 730-741.
- Kain, S. (2008). Entwickler in komplexen Mensch-Maschine-Systemen: Analyse der Antizipationsgüte von Konflikten. In G. f. A. e.V. (Ed.), *54. Kongress der Gesellschaft für Arbeitswissenschaft, Produkt- und Produktions-Ergonomie - Aufgabe für Entwickler und Planer* (pp. 815-818). Dortmund: GFA Press.

- Kain, S., & Nachtwei, J. (2009). Die Rolle von Kontrollvariablen in der Human Factors Forschung: Ein bewährtes Konzept in einem modernen Anwendungsfeld. In B. Krause & P. Metzler (Eds.), *Empirische Evaluationsmethoden* (Band 13, pp. 5-14). Berlin: ZeE Verlag.
- Kaufmann, I., Pornschlegel, H., & Udris, I. (1982). Arbeitsbelastung und Beanspruchung. In L. Zimmermann (Ed.), *Humane Arbeit - Leitfaden für Arbeitnehmer. Band 5: Belastung und Stress bei der Arbeit* (pp. 13-48). Reinbek: Rohwolt.
- Kelsch, J., Flemisch, F. O., Löper, C., Schieben, A., & Schindler, J. (2006). Links oder rechts, schneller oder langsamer? Grundlegende Fragestellungen beim Cognitive Systems Engineering von hochautomatisierter Fahrzeugführung. In M. Grandt & A. Bauch (Eds.), *Cognitive Systems Engineering in der Fahrzeug- und Prozessführung* (pp. 227-240). Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt e.V. (DGLR-Bericht 2006-02/07).
- Kesselring, K. (2009). Entwicklung einer Softwarekomponente zur Systemüberwachung und -kontrolle in einer psychologischen Versuchsumgebung. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Kienast, H., Lindner, P., Weigel, H., Henning, M. J., Krems, J. F., Wanielik, G., et al. (2008). Aufmerksamkeitsausrichtung durch räumlich gerichtete Anzeigen in Fahrerassistenzsystemen. 24. *VW/VDI-Gemeinschaftstagung "Fahrerassistenzsysteme und Integrierte Sicherheit"*, 413-423.
- Kintsch, W., & Ericsson, A. (1996). Die kognitive Funktion des Gedächtnisses. In D. Albert & K. H. Stapf (Eds.), *Kognition, Enzyklopädie der Psychologie* (Themenbereich C: Theorie und Forschung/ Serie II: Kognition/ Band 4: Gedächtnis, pp. 541-601). Göttingen: Hogrefe.
- Knoblich, G., Butterfill, S. A., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In B. H. Ross (Ed.), *Advances in Research and Theory* (pp. 59-101). London: Academic Press.
- Kolbe, M. (2007). Explizite Prozesskoordination von Entscheidungsfindungsgruppen. Unpublished Dissertation. Georg-August-Universität Göttingen.



- Koltko-Rivera, M. E. (2004). Personality Theory and Human Factors Research. In D. Vincenzi, M. Mouloua & P. A. Hancock (Eds.), *Human Performance, Situation Awareness, and Automation: Current research and trends* (Vol. 2, pp. 161-165). Mahwah, New Jersey: Erlbaum.
- Korienek, G., Wrensch, T., & Dechow, D. (2002). *Squeak: A quick trip to objectland*. Boston: Addison-Wesley Verlag.
- Kosjar, N. (2011). Die Gebrauchstauglichkeit des Automaten-GUI im Projekt Arbeitsteilung Entwickler Operateur (ATEO). Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Kosjar, N. (2012). Ein Ereignis-System für das ATEO Automation Framework sowie die Implementierung und Testung von auditiven und visuellen Hinweisen. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Köster, F., & Mehl, K. (2000). Knowledge Discovery in Databases im Kontext der Diagnose und Bewertung von Handlungen in simulierter Realität. In D. P. F. Möller (Ed.), *Frontiers in Simulation - Simulationstechnik. 14. Symposium in Hamburg, September 2000* (pp. 461-466). Ghent: SCS.
- Krause, W. (2000). *Denken und Gedächtnis aus naturwissenschaftlicher Sicht: Vereinfachung, Flexibilität und Invarianz in der menschlichen Informationsverarbeitung*. Göttingen: Hogrefe.
- Krinner, C. (2007a). Developers anticipating users' behavior during design. In M. Kutar (Ed.), *Psychology of Programming Interest Group: Work-in-Progress Workshop* (pp. 1-12). Salford: IRIS.
- Krinner, C. (2007b). How developers anticipate user behavior in the design of assistance systems. In D. Harris (Ed.), *Engineering psychology and cognitive ergonomics* (Vol. 4562, pp. 98-107). Heidelberg: Springer-Verlag.
- Krinner, C. (2009). *Design von Assistenz: Einfluss verschiedener Determinanten auf Assistenzkonzepte von Entwicklern*. Saarbrücken: VDM Verlag Dr. Müller.
- Krinner, C., & Gross, B. (2005). Arbeitsteilung zwischen Entwicklern und Operateuren von Mensch-Maschine-Systemen – eine neue Perspektive auf Funktionsteilung in Mensch-Maschine-Systemen. Unpublished Zusammenfassung des Promotionsvorhabens. Humboldt-Universität zu Berlin, Institut für Psychologie.

- Krinner, C., & Henkel, S. (2007). Entwicklung von Assistenzkonzepten unter verschiedenen ressourcenreichen Bedingungen. In M. Rötting, G. Wozny, A. Klostermann & J. Huss (Eds.), *Prospektive Gestaltung von Mensch-Technik-Interaktion. 7. Berliner Werkstatt Mensch-Maschine-Systeme* (pp. 1-6). Düsseldorf: VDI-Verlag.
- Krinner, C., & Polkehn, K. (2006). Squeak als Simulationswerkzeug in Forschung und Lehre zur MCI. In A. M. Heinecke & H. Paul (Eds.), *Mensch & Computer 2006: Mensch und Computer im Strukturwandel* (pp. 47-72). München: Oldenbourg.
- Kyllonen, P. C., & Christal, P. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, 14(4), 389-433.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174.
- Lang, J. W. B., & Fries, S. (2006). A Revised 10-Item Version of the Achievement Motives Scale - Psychometric Properties in German-Speaking Samples. *Journal of Psychological Assessment*, 22(3), 216-224.
- Laughery, K. R., & Hammond, A. (1999). Overview. In M. S. Wogalter, D. M. Dejoy & K. R. Laughery (Eds.), *Warnings and risk communication* (pp. 3-14). London: Taylor & Francis.
- Lee, J. D., McGehee, D. V., Brown, T. L., & Reyes, M. L. (2002). Collision Warning Timing, Driver Distraction, and Driver Response to Imminent Rear-End Collisions in a High-Fidelity Driving Simulator. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 44(2), 314-334.
- Leonhard, C. (2010). Fenster zum Prozess: Weiterentwicklung eines Operatorsarbeitsplatzes im Projekt ArbeitsTeilung Entwickler Operateur (ATEO). Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Leonhard, C. (2013). Fenster zum Prozess: Weiterentwicklung eines Operatorsarbeitsplatzes im Projekt ArbeitsTeilung Entwickler Operateur (ATEO). Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Lerner, R. (2010). Konzeption und Bewertung eines fahrsituations- und fahrerleistungsadaptiven Warn- und Informationsmanagers. Unpublished Dissertation. Universität der Bundeswehr, Fakultät für Luft- und Raumfahrttechnik, Institut für Arbeitswissenschaft.

- Lerner, N. D., Kotwal, B. M., Lyons, R. D., & Gardner-Bonneau, D. J. (1996). Preliminary human factors guidelines for crash avoidance warning devices. (No. HS-808 342). Washington, D.C.: U.S. Department of Transportation, National Highway Traffic Safety Administration.
- Levi, D., & Slem, C. (1995). Team work in research and development organizations: The characteristics of successful teams. *International Journal of Industrial Ergonomics*, 16(1), 29-42.
- Liebig, T. (1994). Funktionen zur Unterstützung des Piloten bei heutigen Airbus-Flugzeugen. In K.-P. Gärtner (Ed.), *Optimaler Automatisierungsgrad von Mensch-Maschine-Systemen. DGLR-Bericht 94-01*. Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt Lilienthal-Oberth e.V.
- Lindberg, T. (2012). Entwicklung einer ABK-Metapher für gruppierte Fahrerassistenzsysteme. Unpublished Dissertation. Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Mensch-Maschine-Systeme.
- Loosch, E. (1999). *Allgemeine Bewegungslehre*. Wiebelsheim: Limpert Verlag.
- Lüer, G., & Spada, H. (1990). Denken und Problemlösen. In H. Spada (Ed.), *Allgemeine Psychologie* (pp. 189-280). Bern: Huber.
- Maidique, M. A., & Zirger, B. J. (1985). The new product learning cycle. *Research Policy*, 14(6), 299-313.
- Maier, K. (2014). *Multimodaler Warnbaukasten: Entwicklung einer neuen Warnphilosophie für Fahrerassistenzsysteme zur Unterstützung unfallkritischer Bremsszenarien*. Göttingen: Cuvillier Verlag
- Maloney, J. (2000). An Introduction to Morpich: The Squeak User Interface Framework. Retrieved 24. November, 2015, from <http://sdmeta.gforge.inria.fr/FreeBooks/CollectiveNBlueBook/morphich.final.pdf>
- Manzey, D. (2008). Systemgestaltung und Automatisierung. In P. Badke-Schaub, G. Hofinger & K. Lauche (Eds.), *Human Factors: Psychologie sicheren Handelns in Risikobereichen* (pp. 307-324). Heidelberg: Springer Medizin Verlag.
- Marsden, P., & Kirby, M. (2005). Allocation of functions. In N. A. Stanton, A. Hedge, K. Brookhuis, E. Salas & H. Hendrick (Eds.), *Handbook of Human Factors and Ergonomics Methods* (pp. 338-346). London: Taylor & Francis.

- Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for In-Vehicle Information Systems: Annoyance, Urgency, and Appropriateness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(1), 145-157.
- Mayhew, D. J. (1999). *The usability engineering lifecycle: A practitioner's handbook for user interface design*. San Francisco: Morgan Kaufmann Publishers.
- Mayring, P. (2000). Qualitative Inhaltsanalyse. *Forum Qualitative Sozialforschung*. 1(2), Art. 20. Retrieved 12. May, 2008, from <http://www.qualitative-research.net/fqs-texte/2-00/2-00mayring-d.htm>
- Mayring, P. (2007). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (9th ed.). Weinheim und Basel: Beltz Verlag.
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, 52(6), 1258-1265.
- McCrae, R. R., & Costa Jr., P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- McCrae, R. R., & Costa Jr., P. T. (1990). *Personality in adulthood*. New York: Guilford Press.
- McKeown, D., & Isherwood, S. (2007). Mapping Candidate Within-Vehicle Auditory Displays to Their Referents. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(3), 417-428.
- Meyer, B., & Schermuly, C. C. (2011). When beliefs are not enough: Examining the interaction of diversity faultlines, task motivation, and diversity beliefs on team performance. *European Journal of Work and Organizational Psychology*, 21(3), 456-487.
- Meyer, C., Nachtwei, J., & Kain, S. (2009). Empirische Realität versus subjektive Repräsentation: Einfluss von Persönlichkeitsmerkmalen auf die Prozessüberwachung und -führung in Mensch-Maschine-Systemen. *MMI interaktiv*, 12, 1-14.
- Meyer, M., Bekkering, H., Haartsen, R., Stapel, J. C., & Hunnies, S. (2015). The role of action prediction and inhibitory control for joint action coordination in toddlers. *Journal of Experimental Child Psychology*, 139, 203-330.
- Miller, C. A., & Parasuraman, R. (2007). Designing for Flexible Interaction Between Humans and Automation: Delegation Interfaces for Supervisory Control. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 49(1), 57-75.

- Miller, G. A., Galanter, E., & Pribram, K. A. (1960). *Plans and the structure of behavior*. New York: Holt, Rhinehart & Winston.
- Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive Automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied*, 6(1), 44-58.
- Müsseler, J., & Prinz, W. (2002). *Allgemeine Psychologie*. Heidelberg: Spektrum Akademischer Verlag.
- Nachtwei, J. (2008a). *Ein experimenteller Ansatz in der kognitiven Ergonomie: Personenmerkmale als Kodeterminanten der Komplexität einer belebten Mikrowelt*. Saarbrücken: VDM Verlag Dr. Müller.
- Nachtwei, J. (2008b). Fragebogen zur Erfassung von Strategie und Expertise in Experimenten, revidierte Fassung (FESE-R). Unpublished Testhandbuch. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Nachtwei, J. (2011a). Design and Evaluation of a Supervisory Control Lab System for Automation Research: A Theoretical and Empirical Contribution to the Discussion on Funktion Allocation. Unpublished Dissertation. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Psychologie.
- Nachtwei, J. (2011b). A multi-level design approach for a supervisory control master display in human factors experiments. Unpublished Manuscript submitted to Behaviour & Information Technology. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Nachtwei, J., & Kain, S. (2010). Über die ökologische Validität in Untersuchungen zur Mensch-Technik-Interaktion. In B. Krause & P. Metzler (Eds.), *Empirische Evaluationsmethoden* (Band 14, pp. 11-17). Berlin: Zentrum für empirische Evaluationsmethoden.
- Nachtwei, J., & Meyer, C. (2011). Between keyhole and clutter effect: A multi-level evaluation of interface extensions using the ATEO master display (AMD). Unpublished Manuscript submitted to Human Factors. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Nerdinger, F. W., Blickle, G., & Schaper, N. (2011). *Arbeits- und Organisationspsychologie* (2nd ed.). Berlin: Springer Verlag.

- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 152-158). Boston, MA: ACM.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *CHI '90 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 249-256). Seattle. New York: ACM
- Niestroj, N. (2008). Erweiterung des ATEO-Systems zur Komplexitätserhöhung von SAM. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Niestroj, N. (2009). Vernetzung im ATEO Projekt aus inhaltlicher und technischer Sicht. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Niestroj, N. (2014). Entwicklung von Automaten für Mensch-Maschine-Systeme und Ergebnisse ihres Einsatzes. Unpublished Dissertation. Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät, Institut für Psychologie.
- Norman, D. A. (1986). Cognitive engineering. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New Perspectives on Human-computer Interaction* (pp. 31-61). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Norman, D. A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.
- Norman, D. A. (2007). *The Design of Future Things*. New York: Basic Books.
- Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(3), 476-488.
- Ormerod, T. C., & Shepherd, A. (2004). Using task analysis for information requirements specification: the sub-goal template (SGT) method. In D. Diaper & N. A. Stanton (Eds.), *The handbook of task analysis for human-computer interaction* (pp. 347-365). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Pahl, G., Beitz, W., Feldhusen, J., & Grote K.-J. (2007). *Konstruktionslehre. Grundlagen erfolgreicher Produktentwicklung, Methoden und Anwendungen* (7th ed.). Berlin: Springer Verlag.
- Parasuraman, R. (1997). Human use and abuse of automation. In M. Mouloua & J. M. Koonce (Eds.), *Human-Automation interaction: Research and practice* (pp. 42-47). Mahwah, NJ: Lawrence Erlbaum Associates.

- Parasuraman, R. (2000). Designing automation for human use: empirical studies and quantitative models. *Ergonomics*, 43(7), 931-951.
- Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced "complacency". *The International Journal of Aviation Psychology*, 3(1), 1-23.
- Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of Adaptive Task Allocation on Monitoring of Automated Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38(4), 665-679.
- Parasuraman, R., Mouloua, M., Molloy, R., & Hilburn, B. (1993). Adaptive function allocation reduces performance costs of static automation. In J. G. Morrison (Ed.), *The Adaptive Function allocation for Intelligent Cockpits (AFAIC) program: Interim research and guidelines for the application of adaptive automation* (Progress Report. No. NAWCADWAR-93031-60, pp. 37-42). Warminster, PA: Air Vehicle and Crew Systems Technology Department.
- Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39(2), 230-253.
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3), 286-297.
- Parasuraman, R., & Wickens, C. D. (2008). Humans: Still Vital After All These Years of Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50(3), 511-520.
- Peeters, M. A. G., Rutte, C. G., van Tuijl, H. F. J. M., & Reymen, I. M. M. J. (2008). Designing in Teams: Does Personality Matter? *Small Group Research*, 39(4), 438-467.
- Peeters, M. A. G., Tuijl, H. F. J. M. V., Rutte, C. G., & Reymen, I. M. M. J. (2006). Personality and Team Performance: A Meta-Analysis. *European Journal of Personality*, 20(5), 377-396.
- Perrow, C. (1999). *Normal Accidents: Living with High-Risk Technologies* (2nd ed.). Princeton, N.J.: Princeton University Press.

- Pietzcker, F. (2004). *Konstruktion lehren – Wirkung einer konstruktionsmethodischen Ausbildung auf das Konstruieren bei Studenten und Konstrukteuren*. München: Verlag Dr. Hut.
- Pinto, J. K., & Prescott, J. E. (1988). Variations in critical success factors over the stages in the project life cycle. *Journal of Management*, 14(1), 5-18.
- Poling, T., Woehr, D. J., Arciniega, L. M., & Gorman, A. (2006). The Impact of Personality and Value Diversity on Team Performance. *21st SIOP (Society for Industrial and Organizational Psychology, Inc.) Conference*, 1-12.
- Poulton, E. C. (1974). *Tracking Skill and Manual Control*. London: Academic Press.
- Preim, B. (1999). *Entwicklung interaktiver Systeme: Grundlagen, Fallbeispiele und innovative Anwendungsfelder*. Berlin: Springer Verlag.
- Rammstedt, B. (1997). Die deutsche Version des Big Five Inventory (BFI): Übersetzung und Validierung eines Fragebogens zur Erfassung des Fünf-Faktoren-Modells der Persönlichkeit. Unpublished Diplomarbeit. Universität Bielefeld, Fakultät für Psychologie und Sportwissenschaften.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1), 203-212.
- Rammstedt, B., Koch, K., Borg, I., & Reitz, T. (2004). Entwicklung und Validierung einer Kurzskala für die Messung der Big-Five-Persönlichkeitsdimensionen in Umfragen. *ZUMA-Nachrichten*, 55(28), 5-28.
- Rasch, B., Hofmann, W., Frieze, M., & Naumann, E. (2010). Verfahren für Rangdaten. In B. Rasch, W. Hofmann, M. Frieze & E. Naumann (Eds.), *Quantitative Methoden. Band 2: Einführung in die Statistik für Psychologen und Sozialwissenschaftler* (pp. 143-170). Berlin, Heidelberg: Springer-Verlag.
- Rasch, R. H., & Tosi, H. L. (1992). Factors affecting Software Developers' Performance: An Integrated Approach. *MIS Quarterly*, 16(3), 395-413.
- Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York: Elsevier Science Inc.



- Reagans, R., Argote, L., & Brooks, D. (2005). Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science*, 51(6), 869-881.
- Reichart, S. V. (2002). Die Gestaltung des Produktinnovationsprozesses. In M. Reichart (Ed.), *Prozessmanagement mit System* (pp. 97-132). Berlin: Wissenschaftlicher Verlag Berlin.
- Riemer, S., Weißhahn, G., & Hacker, W. (1999). Zur Organisation der Entwurfstätigkeit: Opportunistisches Vorgehen mit geplanten Episoden. *Forschungsbericht der TU Dresden, Institut für Allgemeine Psychologie und Methoden der Psychologie* (Band 68).
- Robinson, G. S., Lee, S. E., & Casali, J. G. (2006). Auditory warnings and displays: issues relating to design and selection. In W. Karwowski (Ed.), *International Encyclopedia of Ergonomics and Human Factors. Volume 1* (2nd ed., pp. 1379-1383). Boca Raton: CRC Press.
- Rocholl, A. (2010). Determinanten der Leistung von Entwicklern von Automaten und Assistenzsystemen zur Überwachung und Führung dynamischer technischer Prozesse. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Rohrmann, B. (1978). Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie*, 9(1), 222-245.
- Röse, K. (2003). Task Analyse. In S. Heinsen & P. Vogt (Eds.), *Usability praktisch umsetzen. Handbuch für Software, Web, Mobile Devices und andere interaktive Produkte* (pp. 98-114). München Wien: Carl Hanser Verlag.
- Rosen, C. C. (2005). The influence of intra-team relationships on the systems development process: A theoretical framework of intra-group dynamics. In P. Romero, J. Good, E. A. Chaparro & S. Bryant (Eds.), *Proceedings of the 17th Workshop of the Psychology of Programming Interest Group* (pp. 30-42). Brighton, UK, Sussex University.
- Royce, W. W. (1987). Managing the Development of Large Software Systems: concepts and techniques. *Proceedings, IEEE WESCON*, 328-338.
- Runco, M. A., & Okuda, S. M. (1988). Problem Discovery, Divergent Thinking, and Creative Process. *Journal of Youth and Adolescence*, 17(3), 211-220.
- Rupp, C. (2006). *Requirements-Engineering und -Management: Professionelle, iterative Anforderungsanalyse für die Praxis*. München: Hanser Verlag.

- Sachse, P., & Furtner, M. (2009). Embodied knowledge in design. In K. Leidlmair (Ed.), *After Cognitivism. A Reassessment of Cognitive Science and Philosophy* (pp. 163-179). Dordrecht: Springer.
- Sachse, P., Hacker, W., Leinert, S., & Riemer, S. (1999). Prototyping als Unterstützungsmöglichkeit des Denkens und Handelns beim Konstruieren. *Zeitschrift für Arbeits- und Organisationspsychologie*, 2, 71-82.
- Sachse, P., & Leinert, S. (1996). Early prototyping II. Ansatz zu einer Modelltaxonomie. *Forschungsbericht der TU Dresden, Institut für Allgemeine Psychologie und Methoden der Psychologie* (Band 37).
- Salas, E. (2008). At the Turn of the 21st Century: Reflections on Our Science. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50(3), 351-353.
- Sarodnick, F., & Brau, H. (2006) Methoden der Usability Evaluation: Wissenschaftliche Grundlagen und praktische Anwendung. Bern: Verlag Hans Huber.
- Sarris, V. (1990). *Methodologische Grundlagen der Experimentalpsychologie 1: Erkenntnisgewinnung und Methodik*. München: Reinhardt.
- Sarris, V., & Reiß, S. (2005). *Kurzer Leitfaden der Experimentalpsychologie*. München: Pearson Studium.
- Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation Surprises. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (2nd ed., pp. 1926-1943). New York: John Wiley & Sons.
- Sauer, J., Wastell, D. G., & Hockey, G. R. J. (2000). A conceptual framework for designing microworlds for complex work domains: a case study of the Cabin Air Management System. *Computers in Human Behavior*, 16(1), 45-58.
- Sauer, J., Wastell, D. G., Hockey, G. R. J., Crawshaw, C. M., & Downing, J. C. (2003). Designing micro-worlds of transportation systems: The computer-aided bridge operation task. *Computers in Human Behavior*, 19(2), 169-183.
- Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Human performance in automated systems: Theory and Applications* (pp. 37-64). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Scherm, M. (1998). Synergie in Gruppen - Mehr als eine Methapher? In E. Ardel-Gattinger, H. Lechner & W. Schlögl (Eds.), *Gruppendynamik. Anspruch und Wirklichkeit der Arbeit in Gruppen* (pp. 61-69). Göttingen: Verlag für Angewandte Psychologie.
- Schermuly, C. C. (2007). Entwicklung und Prüfung eines Messinstruments für Diskussions- und Entscheidungsprozesse: Das Instrument zur Kodierung von Diskussionen (IKD). Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Schermuly, C. C. (2010). Das Instrument zur Kodierung von Diskussionen (IKD): Untersuchung der psychometrischen Qualität und experimenteller Einsatz zur Prüfung des Empowermentkonstrukts. Unpublished Dissertation. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Schermuly, C. C., & Scholl, W. (2010). *Instrument zur Kodierung von Diskussionen (IKD)*. Göttingen: Hogrefe.
- Schermuly, C. C., Schröder, T., Nachtwei, J., & Scholl, W. (2010). Das Instrument zur Kodierung von Diskussionen (IKD): Ein Verfahren zur zeitökonomischen und validen Kodierung von Interaktionen in Organisationen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 54(4), 149-170.
- Schlicksupp, H. (1998). *Ideenfindung*. Würzburg: Vogel Business Media.
- Schmalt, H.-D., & Langens, T. A. (2009). *Motivation* (4th ed.). Stuttgart: Kohlhammer.
- Scholl, W. (2005). Grundprobleme der Teamarbeit und ihre Bewältigung - ein Kausalmodell. In M. Högl & H. G. Gemünden (Eds.), *Management von Teams. Theoretische Konzepte und empirische Befunde* (3rd ed., pp. 33-66). Wiesbaden: Gabler.
- Schütze, M. (2003). Die frühen Phasen des konstruktiven Entwerfens – Unterstützungspotential verschiedenartiger Arbeitsmittel. Unpublished Dissertation. Technische Universität Dresden, Fakultät Mathematik und Naturwissenschaften.
- Schwarz, G. (2008). *Führen mit Humor: Ein gruppendynamisches Erfolgsrezept* (2nd ed.). Wiesbaden: Gabler Verlag.
- Schwarz, H. (2009). Fenster zum Prozess: ein Operatorsarbeitsplatz zur Überwachung und Kontrolle von kooperativem Tracking. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.

- Schwarz, J. (2007). *Code of Practice für die Entwicklung, Validierung und Markteinführung von künftigen Fahrerassistenzsystemen (ADAS). Rechtliche Rahmenbedingungen zur Einführung von Fahrerassistenzsystemen*. Bergisch Gladbach: BAST.
- Scott, J. J., & Gray, R. (2008). A Comparison of Tactile, Visual, and Auditory Warnings for Rear-End Collision Prevention in Simulated Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(2), 264-275.
- Sedlmeier, P., & Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie*. München: Pearson Studium.
- Seid, A. (2012a). Entwicklung und Konfiguration von visuellen Anzeigen für die Objektsteuerung im Projekt ArbeitsTeilung Entwickler-Operator (ATEO). Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Seid, A. (2012b). Erweitern der Log-Datei und der Analyse von Log-Dateien im ATEO Projekt. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Seier, E. (2002). Comparison of Tests of Univariate Normality. Retrieved 08. August, 2015, from <http://interstat.statjournals.net/YEAR/2002/articles/0201001.pdf>
- Seifert, R. (1994). Ergonomische Grundlagen der Automatisierung von Funktionen in MMS. In K.-P. Gärtner (Ed.), *Optimaler Automatisierungsgrad von Mensch-Maschine-Systemen* (DGLR-Bericht 94-01, pp. 33-40). Bonn: Deutsche Gesellschaft für Luft- und Raumfahrt Lilienthal - Oberth e.V.
- Selcon, S. J. (1999). Using redundancy in the design of time-critical warnings: a theory-driven approach. In N. A. Stanton & J. Edworthy (Eds.), *Human factors in auditory warnings* (pp. 201-224). Aldershot: Ashgate.
- Sfetsos, P., Stamelos, I., Angelis, L., & Deligiannis, I. (2009). An experimental investigation of personality types impact on pair effectiveness in pair programming. *Empir Software Eng*, 14, 187-226.
- Sharit, J. (1984). Human supervisory control of a flexible manufacturing system: An exploratory investigation. Unpublished Dissertation. West Lafayette, IN: Purdue University.
- Sheridan, T. B. (1988). Task allocation and supervisory control. In M. Helander (Ed.), *Handbook of Human-Computer Interaction* (pp. 159-173). Amsterdam: Elsevier Science Publishers.

- Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control* (2nd ed.). Cambridge, MA: MIT Press.
- Sheridan, T. B. (2000). Function allocation: algorithm, alchemy or apostasy? *International Journal of Human-Computer Studies*, 52(2), 203-216.
- Sheridan, T. B. (2002). *Humans and Automation: System Design and Research Issues*. New York: Wiley.
- Sheridan, T. B., & Parasuraman, R. (2006). Human-automation interaction. *Reviews of Human Factors and Ergonomics*, 1(1), 89-129.
- Sheridan, T. B., & Verplank, W. L. (1978). *Human and computer control of undersea teleoperators (Man-Machine Systems Laboratory Report)*. Cambridge (MA): Massachusetts Institute of Technology.
- Sicinski, A. (1995). On Designer's Responsibility. In A. Collen & W. W. Gasparski (Eds.), *Design and systems: General applications of methodology* (Vol. 3, pp. 403-412). New Brunswick, NJ: Transaction Publisher.
- Simon, H. A. (1995). Problem forming, problem finding, and problem solving in design. In A. Collen & W. W. Gasparski (Eds.), *Design and systems: General applications of Methodology* (pp. 245-257). New Brunswick, NJ: Transaction Publisher.
- Skjerve, A. B. M., & Skraaning Jr., G. (2004). The quality of human-automation cooperation in human-system interface for nuclear power plants. *International Journal of Human-Computer Studies*, 61(5), 649-677.
- Smith, G. F., & Browne, G. J. (1993). Conceptual Foundations of Design Problem Solving. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(5), 1209-1219.
- Smith, P., McCoy, E., & Layton, C. (1997). Brittleness in the Design of Cooperative Problem-Solving Systems: The Effects on User Performance. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*, 27(3), 360-371.
- Sommerville, I. (2007). *Software Engineering* (8th ed.). Boston: Addison-Wesley Verlag.
- Sonnentag, S. (1994). Streß in SE-Projekten. In F.C. Brodbeck & M. Frese (Eds.), *Produktivität und Qualität in Software-Projekten* (pp. 71-85). Wien: R. Oldenbourg Verlag.

- Stade, M. J. C. (2012). Verhaltensindikatoren von Operateuren zur Evaluation der Prozessüberwachungs- und Führungsgüte am Beispiel einer simulierten Leitwarte im ATEO Projekt. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Psychologie.
- Stade, M. J. C., Meyer, C., Niestroj, N., & Nachtwei, J. (2011). (Not) everybody's darling: Value and prospects of multiple linear regression analysis and assumption checking. In B. Krause, R. Beyer & G. Kaul (Eds.), *Empirische Evaluationsmethoden* (Band 15, pp. S. 17-37). Berlin: Zentrum für empirische Evaluationsmethoden.
- Stanton, N. (2006). Hierarchical task analysis: Developments, applications, and extensions. *Applied Ergonomics*, 37(1), 55-79.
- Stanton, N., & Edworthy, J. (1999). Auditory warning affordances. In N. Stanton & J. Edworthy (Eds.), *Human factors in auditory warnings* (pp. 113-127). Aldershot: Ashgate.
- Steinheider, B. (2001). Supporting the co-operation of R&D-Teams in the product development process. *Proceedings of the 5th Conference on Engineering Design and Automation*, 5-8 August 2001, Las Vegas, Nevada.
- Stevens, C. J., Brennan, D., & Parker, S. (2004). Simultaneous manipulation of parameters of auditory icons to convey direction, size, and distance: Effects on recognition and interpretation. *Proceedings of the ICAD '04 - Tenth Meeting of the International Conference on Auditory Display*, 6-9 July 2004, Sydney, Australia, 1-7.
- Stevens, C. J., Brennan, D., Petocz, A., & Howell, C. (2009). Designing informative warning signals: Effects of indicator type, modality, and task demand on recognition speed and accuracy. *Advances in Cognitive Psychology*, 5, 84-90.
- Susman, G. I., & Ray, J. M. (1999). Test of a model of organizational contributors to product development team effectiveness. *Journal of Engineering and Technology Management*, 16(3-4), 223-245.
- Swain, A. D., & Guttman, H. E. (1983). *Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final Report*. Washington, DC: U.S. Nuclear Regulatory Commission.
- Thoma, S. (2010). Mensch-Maschine-Interaktionskonzepte für Fahrerassistenzsysteme im Kreuzungsbereich. Unpublished Dissertation. Technische Universität München, Institut für Informatik.

- Timpe, K.-P. (1998). Unterstützungssysteme als interdisziplinäre Herausforderung: Einführung in die Tagung "Wohin führen Unterstützungssysteme? In H.-P. Willumeit & H. Kolrep (Eds.), *Wohin führen Unterstützungssysteme? Entscheidungshilfe und Assistenz in Mensch-Maschine-Systemen*, 2. Berliner Werkstatt Mensch-Maschine-Systeme, 7.-9. Oktober 1997 (ZMMS-Spektrum Band 5, pp. 1-20). Sinzheim: Pro Universitate Verlag.
- Timpe, K.-P., & Kolrep, H. (2000). Das Mensch-Maschine-System als interdisziplinärer Gegenstand. In K.-P. Timpe, T. Jürgensohn & H. Kolrep (Eds.), *Mensch-Maschine-Systemtechnik: Konzepte, Modellierung, Gestaltung, Evaluation* (pp. 9-40). Düsseldorf: Symposion Publishing.
- van Knippenberg, D., & Schippers, M. (2007). Work group diversity. *Annual Review of Psychology*, 58, 515-541.
- VDI (1993). *VDI Richtlinien 2221: Methodik zum Entwickeln und Konstruieren technischer Systeme und Produkte*. Düsseldorf: VDI-Verlag.
- Verworn, B. (2005). *Die frühen Phasen der Produktentwicklung. Eine empirische Analyse in der Mess-, Steuer- und Regelungstechnik*. Wiesbaden: Deutscher Universitätsverlag.
- Verworn, B., & Herstatt, C. (2007). Bedeutung und Charakteristika der frühen Phasen des Innovationsprozesses. In C. Herstatt & B. Verworn (Eds.), *Management der frühen Innovationsphasen: Grundlagen – Methoden – Neue Ansätze* (2nd ed., pp. 3-19). Wiesbaden: Gabler.
- Visser, W. (1994). Organisation of design activities: opportunistic, with hierarchical episodes. *Interacting with Computers*, 6(3), 239-274.
- Vöhringer-Kuhnt, T. (2008). CarUSE: Entwicklung eines Beschreibungsinstrumentes für Human-Machine Interface unter besonderer Berücksichtigung von Fahrerinformations- und Fahrerassistenzsystemen. Unpublished Exposé zur Dissertation. Fakultät V - Verkehrs- und Maschinensysteme der Technischen Universität Berlin, Institut für Psychologie und Arbeitswissenschaft, Fachgebiet Mensch-Maschine-Systeme.
- Vöhringer-Kuhnt, T. (2010). CarUSE-MI: Die Entwicklung eines Bewertungsinstrumentariums für die Mensch-Maschine-Schnittstelle von Fahrerinformationssystemen. Unpublished Dissertation. Fakultät V - Verkehrs- und Maschinensysteme der Technischen Universität Berlin.

- Volpert, W. (1987). Psychische Regulation von Arbeitstätigkeiten. In U. Kleinbeck & J. Rutenfranz (Eds.), *Enzyklopädie der Psychologie, Themenbereich D, Serie III, Band 1: Arbeitspsychologie* (pp. 1-42). Göttingen: Hogrefe.
- von Bernstorff, C. (2014). Effekte automatischer Unterstützung auf die Prozessüberwachungs- und -führungsleistung von Operateuren: Theoretischer und empirischer Beitrag auf dem Weg zur kooperativen Automation. Unpublished Dissertation. Humboldt-Universität zu Berlin, Lebenswissenschaftliche Fakultät, Institut für Psychologie.
- von der Weth, R. (2001). *Management der Komplexität: ressourcenorientiertes Handeln in der Praxis*. Bern: Huber.
- Waern, Y., & Canas, J. J. (2003). Microworld task environments for conducting research on command and control. *Cognition, Technology & Work*, 5(3), 181-182.
- Wagner, A. (2012a). Restrukturierung der Architektur einer weichen Echtzeitanwendung im Kontext von Variabilitätsmanagement. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Wagner, A. (2012b). Variabilitätsmanagement im Kontext von software-basierten Experimentalsystemen. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Walker, D. A. (2003). JMASM9: Converting Kendall's Tau For Correlational Or Meta-Analytic Analyses. *Journal of Modern Applied Statistical Methods*, 2(2), 525-530.
- Wandke, H. (2005). Assistance in human-machine interaction: A conceptual framework and a proposal for a taxonomy. *Theoretical Issues in Ergonomics Science*, 6(2), 129-155.
- Wandke, H., & Nachtwei, J. (2008). The different human factor in automation: the developer behind vs. the operator in action. In D. de Waard, F. O. Flemisch, B. Lorenz, H. Oberheid & K. A. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 493-502). Maastricht, the Netherlands: Shaker Publishing.
- Wandke, H., Oed, R., Metzker, E., van Ballegooy, M., & Nitschke, J. (2001). Die Entwicklung von User Interfaces als arbeitswissenschaftlicher Prozess und seine Unterstützung durch Software-Werkzeuge. *Zeitschrift für Arbeitswissenschaft*, 55(2), 79-93.
- Wandke, H., & Wetzenstein, E. (2004). Assistenz für Benutzer: Den Umgang mit Technik erleichtern. *VDI-Reihe 22: Mensch-Maschine-Systeme*.



- Wandmacher, J. (1993) Software-Ergonomie. In H. Balzert (Series Ed.), *Mensch Computer Kommunikation - Grundwissen 2*. Berlin: Walter de Gruyter & Co.
- Washington, D. (2013). Entwicklung und Implementation eines Prozesses zur Produktableitung aus einer Produktlinie im Kontext von software-basierten Experimentalsystemen. Unpublished Bachelorarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Weck, M. (2006). Knowledge Creation and Exploitation in Collaborative R&D Projects: Lessons Learned on Success Factors. *Knowledge and Process Management*, 13(4), 252-263.
- Weidner-Kim, H. (2014). Konzeption und Implementation einer optimalen Rückführung in ATEO SAM. Unpublished Studienarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Weidner-Kim, H. (2015). Konzeption und Implementation einer verbesserten optimalen Rückführung in ATEO SAM. Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Weiner, B. (1994). *Motivationspsychologie* (3rd ed.). Weinheim: Beltz, Psychologie Verlag Union.
- Weißhahn, G., & Rönsch, T. (2002). Unterstützungen von Entscheidungen - computergestützte Entscheidungssysteme warum und wie? In W. Hacker (Ed.), *Denken in der Produktentwicklung. Psychologische Unterstützung der frühen Phasen* (pp. 111-128). Zürich: vdf Hochschulverlag AG & Rainer Hampp.
- Weller, G., & Schlag, B. (2002). Kriterien zur Beurteilung von Fahrerassistenzsystemen. In 38. *BDP-Kongress für Verkehrspsychologie* (pp. 1-10). Regensburg.
- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159-177.
- Wickens, C. D. (2007). Attention to Attention and its Applications: A Concluding View. In A. F. Kramer, D. A. Wiegmann & A. Kirlik (Eds.), *Attention: From Theory to Practice* (pp. 239-249). New York: Oxford University Press.
- Wickens, C. D., Lee, J. D., Liu, Y., & Gordon-Becker, S. (2004). *An Introduction to Human Factors Engineering* (2nd ed.). Upper Saddle River (NJ): Pearson Education.

- Wickert, A. (2013). Klassifizierung und Entwicklung von Automaten für Eingriffe in der Socially Augmented Microworld (SAM). Unpublished Diplomarbeit. Humboldt-Universität zu Berlin, Institut für Informatik.
- Wiczorek, R., & Zinn, J. (2008). Berücksichtigung von HF-Aspekten im Anlagendesign - Eine Methode zur Funktionsallokation. Unpublished Kursarbeit.
- Wiener, E. L., & Curry, R. E. (1980). Flight-deck automation: promises and problems. *Ergonomics*, 23(10), 995-1011.
- Wilson, T., Butler, W., McGehee, D., & Dingus, T. A. (1997). Forward-Looking Collision Warning System Performance Guidelines. *SAE Technical Paper 970456 (DOI: 10.4271/970456)*.
- Winkelmann, C., & Hacker, W. (2006). Design Problem Solving: Wovon hängen Lösungsgüteverbesserungen durch eine Frage-Antwort-Technik ab? *Zeitschrift für Psychologie*, 214(2), 73-86.
- Winner, H., Hakuli, S., & Wolf, G. (2009). Einleitung. In H. Winner, S. Hakuli & G. Wolf (Eds.), *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort* (2nd ed., pp. 1-2). Wiesbaden: Vieweg + Teubner.
- Wright, M. C., & Kaber, D. B. (2005). Effects of Automation of Information-Processing Functions on Teamwork. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 47(1), 50-66.
- Wübbelmann, K. (2001). *Management Audit. Unternehmenskontext, Teams und Managerleistung systematisch analysieren*. Wiesbaden: Gabler.
- Zijlstra, F. R. H. (1993). Efficiency in work behaviour: a design approach for modern tools. Unpublished Dissertation. Delft University of Technology.
- Zijlstra, F. R. H., & van Doorn, L. (1985). *The construction of a scale to measure subjective effort* (Technical Report). Delft, The Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.
- Zöfel, P. (2003). *Statistik für Psychologen: Im Klartext*. München: Pearson Studium.

## **Erklärung der selbstständigen Erarbeitung der Dissertation**

Ich versichere hiermit, dass ich die Dissertationsschrift mit dem Titel:

„Entwickler von komplexen Mensch-Maschine-Systemen:

Analyse des Einflusses von deren Leistungsdeterminanten

auf den Entwicklungsprozess und das -ergebnis“

in allen Teilen selbstständig verfasst und keine unerlaubten Hilfsmittel benutzt habe.

Die Stellen, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen sind, habe ich in jedem einzelnen Fall durch Angabe der Quelle, auch der benutzten Sekundärliteratur, als Entlehnung kenntlich gemacht.

Ich habe mich nicht anderwärts um einen Dokortitel beworben und besitze noch keinen Doktorgrad im Promotionsfach Psychologie. Die dem Verfahren zu Grunde liegende Promotionsordnung ist mir bekannt.

Berlin, 15.01.2016

---

Saskia Kain